Functional Characterization of Sensor Integration in Distributed Sensor Networks†

Lakshman Prasad and S.S.Iyengar
Department of Computer Science
Louisiana State University
Baton Rouge, LA 70803.

R.L.Kashyap and R.N. Madan
Dept. of Elec. Eng.
Purdue University
West Lafayette, IN 47907

R.N. Madan
Division of Electronics
Office of Naval Research
Arlington, VA 22217-5000

Abstract

Fault-tolerance is an important issue in network design because sensor networks must function in a dynamic, uncertain world. In this paper, we propose a functional characterization of the fault-tolerant integration of abstract interval estimates. This model provides a test bed for a general framework which we hope to develop to address the general problem of fault-tolerant integration of abstract sensor estimates. We further propose a scheme for narrowing the width of the sensor output in a specific failure model and give it a functional representation.

The main distinguishing feature of our model over the original Marzullo's model is in reducing the width of the output interval estimate significantly in most cases where the number of sensors involved is large.

1.0 INTRODUCTION

In recent years, the increasing sophistication of surveillance systems and tracking mechanisms has generated a great deal of interest in the development of new computational structures and strategies for detecting and tracking multiple targets, using data from many sensors.

The design of spatially distributed target-detection-and-tracking systems involves the integration of solutions obtained by solving subproblems in data-association, hypothesis-testing, data-fusion, etc.This must include the cooperative solution of problems by a decentralized and loosely coupled collection of processors, each of which integrates information received from a cluster of spatially distributed sensors into a manageable and reliable output for further integration at a higher level. Integration of information at the sensor level requires techniques to be developed to abstractly represent and integrate sensor information. Further these techniques have to be robust in the sense that even if some of the sensors are faulty, the integrated output should still be reliable. For details on multi sensor integration and fusion in intelligent systems, see [BiBr 90, LuKa 89, HAMD 87, LuLi 88, Duwh 88, Zhen 89].

The aim of this paper is to present a fault-tolerant computational model for sensor integration in *Distributed Sensor Networks.*

A *Distributed Sensor Network* (DSN) consists of spatially distributed sensors which detect and measure a certain phenomenon via its changing parameters. These readings are sent at regular intervals of time to processing units which

† Research sponsored in part by an LEQST grant from LSU Board of Regents and Office of Naval Research. This project is also funded by JPL-Cal Tech grant (Jan 1991). Also, this work was supported in part by the IST Program of the SDIO, monitored by the ONR N00014-85K-0611 and by the NSF Engineering Center Grant CDR-8803017.

integrate the readings from clusters of sensors and give outputs whose nature is much the same as the inputs of the sensors. Outputs from processors representing clusters of sensors are later integrated to get a complete picture of the spatially distributed phenomenon. However, before integration is performed at the processor level, it is necessary to have reliable estimates at each processor. Each sensor in a cluster measures the same set of parameters. It is possible that some of these sensors are faulty. Hence it is desirable to make use of the redundancy of the readings in the cluster to

obtain a correct estimate of the parameter being read. In short, a *fault-tolerant* technique of sensor integration to obtain the correct estimate is sought.

1.1 Scope of This Paper

This paper has two objectives: The first objective is to propose a functional characterization of fault-tolerant integration of abstract interval estimates considered by Marzullo [Marz 89]. The second objective is to propose a modified computational scheme of integration carried out by Marzullo [Marz 89] in the case when the number of sensors is large, wherein it is possible to improve the accuracy of the integraed output.

The main distinguishing feature of our model over the original Marzullo's model is in reducing the width of the output interval estimate significantly in most cases where the number of sensors involved is large.

Elsewhere we intend to generalize Marzullo's approach to the cases when the sensor outputs are subsets of an abstract parameter space. The functional characterization of the fault-tolerant integration of abstract interval estimates described in this paper hints at an abstract framework. We hope to develop for addressing the general problem of fault-tolerant integration of sensor outputs.

1.2 Organization of the Paper

In section 2, we describe Marzullo's work on sensor integration and other related work. Our abstract model functional characterization is detailed in section 3 and is an extension of the model proposed by Marzullo. In section 4, we motivate the need for a new failure model and present the information integration algorithm with a specific example. Finally, we close the paper with concluding remarks and future directions this reach would take.

2.0 RELATED WORK

Marzullo [Marz 89] considers the case of a processor receiving input from several sensors whose outputs are connected intervals. He gives a fault-tolerant integration algorithm which takes as input the intervals representing the sensors and gives as output of the processor a connected interval representing the sensor values. More precisely: Let there be n sensors, each of which yields an interval as its output. there sensors measure a certain physical value and their intervals contain the physical value unless they happen to be faulty sensors.

Thus, a correct sensor is one which contains the actual physical value in its interval. Any two correct sensors must overlap since they both contain the physical value being measured.

Marzullo considers the case when almost f sensors are faulty and gives an algorithm which yields a connected interval as the output of the processor, containing the physical value.

If at most f of the n sensors are faulty, then it follows that at least n-f sensors are correct. Marzullo considers all possible nonempty (n-f)-intersections of the n-sensors. A sensor which does not belong to any of the (n-f)-cliques is faulty since a correct sensor overlaps with at least (n-f-1) other correct sensors. One and only one of the (n-f)-intersections contains the physical value. Since it is not possible to decide which intersection has the physical value (which is as yet unknown to us) and since the processor output is required to be a connected interval, the smallest connected interval containing all the (n-f)-intersections is taken to be the output of the processor. It is easy to see that it contains the actual physical value. The wider this interval is, the lesser the accuracy of the processor output. Marzullo proves the existence of bounds for the width of this interval in terms of f.

The example described below provide a description of integration process.

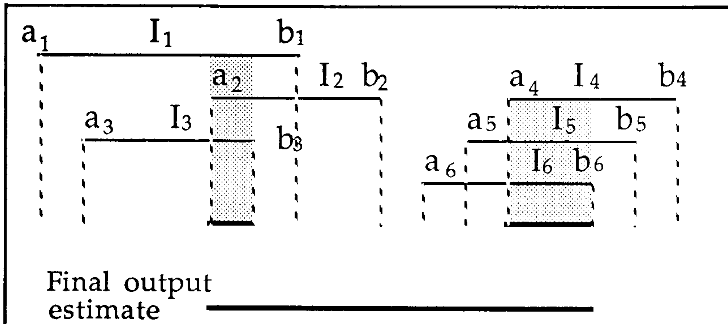

Figure 1. Integration Interval estimates

$(a_1 < a_3 < a_2 < b_3 < b_1 < b_2 < a_6 < a_5 < a_4 < b_6 < b_5 < b_4)$

In the above figure, we have the intervals $I_j = [a_j, b_j]$ $1 \leq j \leq 6$. Overlapping one another according to the strict chain of inequalities given above. Here $f = 3$ and $n = 6$. So, taking all possible $(n - f)$ intersections gives us the intervals $[a_2, b_3]$ and $[a_4, b_6]$. Then enclosing these intervals in the smallest possible connected interval, we have the integrated output interval I_p given by $I_p = [a_2, b_6]$.

Kashyap et. al, [ChKa 89a, ChKa 89b, HaKa 90] have worked with belief intervals to solve decision problems in Expert Systems using Fuzzy Set Theoretic techniques. The belief intervals are however subsets of the closed interval [0,1] and are attached to statements as valuations to aid in Fuzzy reasoning. The intervals considered in this paper are connected subsets of the real line.

More specifically, the problem addressed in our paper is that of obtaining reliable intervals in which the correct physical value being measured lies by taking intersections of appropriate intervals representing sensor outputs rather than obtaining rules of combination of uncertainty intervals of logical statements and syllogisms to obtain uncertainty intervals of compound statements.

The main thrust in our paper is in the derivation of computational schemes for narrowing the width of the processor output in a specific failure model and give it a functional representation.

2.1 *Interval Representation of Sensor Readings*

A sensor reads a physical variable and gives a number as its output. However a sensor is prone to inaccuracies and there may be some uncertainty in the value of its output. The simplest modeling of this is achieved by looking upon sensor outputs as connected intervals on the real line rather than as points. The actual value representing the physical variable being measured is taken to be contained in the interval associated with the sensor if the sensor is not faulty. No assumptions are made about the width of these intervals or their position on the real line. Thus each sensor value is represented by an interval estimate. We make this notion precise in the following definitions that are useful in characterizing one model of sensor integration.

Definition 1 : An *abstract* sensor is a sensor which reads a physical parameter and gives out an abstract interval estimate I_s, which is a bounded and connected subset of the real line **R**.

Definition 2 : A *correct* sensor is an abstract sensor whose interval estimate contains the actual value of the parameter being measured.

Definition 3 : Let sensors $s_1,\ldots,s_n$ feed into a processor P. Let the abstract interval estimate of s_j be I_j $1 \leq j \leq n$, where I_j is the closed interval $[a_j,b_j]$ with end-points a_j and b_j . Define the *characteristic function* χ_j of the j^{th}

sensor s_j $1 \le j \le n$ as follows :

$$\chi_j : \mathbf{R} \rightarrow \{0,1\}$$

$$\chi_j(x) = \begin{cases} 1 \ \forall \ x \in I_j \\ 0 \ \forall \ x \notin I_j \end{cases} \quad \forall \ 1 \le j \le n.$$

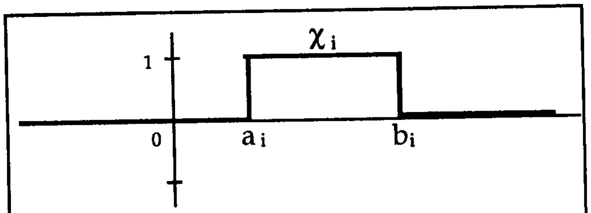

Figure 2. Representation of intervals

The next section addresses the question of how the abstract sensors or abstract estimates are combined to yield new abstract estimates.

3.0 THE PROBLEM OF FAULT-TOLERANT SENSOR INTEGRATION

The problem of fault-tolerant sensor integration is the integration of the I_j $(1 \le j \le n)$, to obtain an abstract interval estimate $I_p = [a_p, b_p]$ which is a 'reliable' and 'fairly accurate' estimate of the region in which the physical sensor value lies.This integration should be fault-tolerant in that its reliability should not be severely affected by some of the sensors being faulty. In other words, we seek to obtain a functional relationship between the characteristic function χ_p of $I_p = [a_p, b_p]$ and the χ_j $1 \le j \le n$: $\chi_p(x) = f(\chi_1(x), ..., \chi_n(x))$ such that $\chi_p^{-1}(1)$ is a fault-tolerant interval estimate of the physical value being measured.

We now go about obtaining a functional representation of the integrated output estimate under the integration scheme of Marzullo. In order to do this we need to introduce a few relevant operations and functions. The following facts provide such operations for our integration problem:

Fact 1: If f(x) is a real-valued function, define $\|f\| = \sup\{|f(x)| \mid x \in \mathbf{R}\}$ (norm of f) . That is, $\|f\|$ is the smallest real number α such that $f(x) \le \alpha \ \forall \ x \in \mathbf{R}$.

Fact 2: If f(x) is a real valued function define Supp $f = \{x \mid f(x) \neq 0\}$ (support of f).

Fact 3: Let $O(x) = \sum_{j=1}^{n} \chi_j(x)$ to be the 'overlap function'. For each $x \in \mathbf{R}$, O(x) gives the number of intervals in which x lies or the number of intervals overlapping at the point x.

Remark 1: The integer $\|\chi_i O(x)\|$ gives the maximum number of intervals in which any $x \in f_i$.

Proof: Indeed, O(x) is the number of intervals overlapping at the point x. Multiplying by $\chi_i(x)$ restricts)O(x) to the interval I_i

$\|\chi_i O\|$ therefore gives the maximun number of intervals overlapping at any $x \in f_i$.

Then Ii belongs to $\|\chi_i O\|$-clique. Where, by a n-clique we mean a group of n intervals having a common intersection.

■

If $\chi_i(x)$ and $\chi_j(x)$ are characteristic functions of intervals I_i and I_j then the characteristic function of the interval $I_i \cap I_j$ denoted $\chi_{I_i \cap I_j}(x)$ is given by the product $\chi_i(x)$ $\chi_j(x)$.

If $\chi_i(x)$ is the characteristic function of I_i, then the characteristic function of I_i^c (the complement set of I_i) is given by $(1 - \chi_i(x))$.

Thus if $I_1, ..., I_n$ are intervals with characteristic functions $\chi_1, ..., \chi_n$, then the characteristic function of their union $\chi_{\cup I_i}(x)$ is given by

$$\chi_{\cup Ii}(x) = 1 - \prod_{i=1}^{n}(1-\chi_i(x)).$$

Marzullo[Marz 89] assumes that there are at most f faulty sensors among n sensors, and considers the intersections of (n-f) or more sensors as the regions in which the correct physical value lies. An interval which does not participate in any (n-f)-intersection is taken to be the estimate of a faulty sensor. The output is the smallest interval which contains all (n-f) or more intersections.

3.1. *Computational Characterization*

Remark 2: If at most f sensors are faulty, then we need to consider only those I_i's for which $\|\chi_i O\| \geq$ (n-f). Thus the characteristic function of the set of all points lying in (n-f) or more intersections of the intervals I_j $(1 \leq j \leq n)$ is given by :

$$S(x) = 1 - \prod_{j=1}^{n}(1 - \chi_{[n-f,\infty)}(\|\chi_i O\|)\chi_j(x))$$

Where $\chi_{[n-f,\infty)}$ is the characteristic function of the interval $[n-f,\infty)$.

Proof: Indeed $\chi_{[n-f,\infty)}(\|\chi_i O\|) = 1$ iff $\|\chi_i O\| \geq n-f$

i.e. iff I_j has at least n-f-1 other intervals intersecting it at some point in it.

So $\chi_{[n-f,\infty)}(\|\chi_i O\|)\chi_j(x) = \chi_j(x)$ if and only if I_j is as described above. So that S(x) is the support of the characteristic function of all those points which belong to (n-f) or more intervals' intersection. ■

Now the correct physical value belongs to Supp(S(x)), i.e; to one of the intervals constituting it. Marzullo proposes the smallest connected interval containing Supp(S(x)) for the integrated output.

More precisely the output interval estimate Ip is given by :

$$I_p = [\,\mathrm{Min}\,\{x \mid S(x) = 1\},\ \mathrm{Max}\,\{x \mid S(x) = 1\}]$$

The above integration technique does indeed give a connected Interval within which the actual physical values lies. It however includes points which do not belong to the intervals constituting Supp(S(x)).

Furthermore, if the intervals constituting Supp(S(x)) are widely scattered over $\bigcup_{i=1}^{k} I_j$, then I_j suffers from inaccuracy since it tends to be very broad.

4.0. A NEW FAILURE MODEL WITH SHARPER OUTPUT INTERVAL ESTIMATES

We propose a failure model in which it is possible to choose in most cases a subset of Supp(S(x)) as the region of correct sensor value instead of the whole of Ip as defined above. A sensor may fail *wildly*, in which case there is no correlation between the actual physical value being measured and the interval estimate of the faulty sensor. On the other hand, a sensor may fail *tamely*, in which case although the faulty sensor's interval estimate does not contain the actual physical value, the interval estimate lies significantly close to the value in a certain sense. For example, mechanical vibrations may induce a tame fault in dials and meters by shifting the needless fluctuations to a region which does not contain the correct value but lies close to it. Since we do not know the actual physical value, we cannot detect the tameness of a fault directly. However tamely faulty sensor estimates tend to overlap with correct sensor estimates because of their proximity to the actual physical value. We consider the case when the number of sensors to be integrated is very large and assume that most of the faulty sensors are tamely faulty. In this case, we observe that correct sensors have a relatively larger number of intervals overlapping with them as compared to undetected faulty sensors participating in the (n-f)-intersections, since tamely faulty sensors overlap with correct sensor estimates.Thus the number of sensor estimates overlapping with a given sensor estimate is a good index of its correctness. We make use of this observation to narrow our output interval estimate, namely Ip.

Let $\mathrm{Supp}(S(x)) = \bigcup_{i=1}^{k} L_i$ where $L_i = [\alpha_i, \beta_i]$ with $\beta_i < \alpha_{i+1}\ \forall\ 1 \le i \le k-1$. We now perform an evaluation of the L_i's in order to attach a weight to each of them and choose those L_i's with maximum weight to be the intervals which have a high likelihood of containing the correct physical value. We then again enclose these L_i's of maximum weight by the smallest possible interval and take it to be the output estimate.

__Remark 3__: Let $\chi_{L_i}(x)$ be the characteristic function of L_i. Then we can define the *popularity* of the j^{th} sensor to be the number $P_j = \sum_{k=1}^{n} \|\chi_k \chi_j\| - 1$. P_j gives the number of sensor intervals overlapping with the j^{th} sensor interval.

__Proof__: Indeed $\|\chi_k \chi_j\| = \begin{cases} 1 & \text{if } I_k \cap I_j \neq \phi \\ 0 & \text{if } I_k \cap I_j = \phi \end{cases}$

Thus $\sum_{k=1}^{n} \|\chi_k \chi_j\| - 1$ gives the number of intervals (apart from I_j) intersecting with I_j.

4.1 Narrowing of the Output Interval Estimate Width

We would like to take the sum of the popularities of all sensors involved in the formation of L_i, and call it the *reliability* r_i of L_i.

Consider the set function

$W(L_i) = \sum_{j=1}^{n} \|\chi_{L_i} \chi_j\| P_j$, $1 \le i \le k$ defined on each L_i.
$W(L_i)$ gives the sum of the number of intervals overlapping with each sensor estimate in the (n-f)-or-more clique L_i.

i.e., $r_i = W(L_i)\ \forall\ 1 \le i \le k$.

Let $r = \max\{r_i \mid 1 \le i \le k\}$, $m = \min\{i \mid r_i = r\}$ and $M = \max\{i \mid r_i = r\}$. Consider the interval $[\alpha_m, \beta_M]$. We take $I_p^* = [\alpha_m, \beta_M]$ as the integrated output estimate.

It is clear that $\beta_M - \alpha_m = |I_p^*| \le |I_p|$, where $|I|$ is the width of the interval I. Thus in our failure model we have in general a way of narrowing the output estimate I_p to I_p^*. However if the number of wildly faulty sensors are as many as the tamely faulty ones, and if they happen to cluster somewhere else on I_p, then it is possible that $|I_p^*| = |I_p|$. Thus the worst case for I_p^* is I_p. The chances that wildly faulty sensors mimic the clustering behavior of tamely faulty sensors are remote. Also if the number of sensors is very small, it is possible that $|I_p^*| = |I_p|$.

For example, consider the case of three input sensor estimates $I_1=[2.4,3.2]$, $I_2=[2.9,4]$, $I_3=[3.6,5]$. In this case $I_p=[2.9,4]$. Here $L_1=[2.9,3.2]$ and $L_2=[3.6,4]$, but they both have the same reliability.Hence $I_p^*=I_p$ here.

4.2 The Algorithm

We now present the algorithm as follows:

Algorithm :

Input : Intervals $I_1, I_2, \ldots, I_n$, f.

Output: Integrated output estimate.

begin

1. **Take** all (n-f)-intersections of the intervals to yield Intervals $L_1, \ldots, L_k$, each of which is an (n-f)-intersection : { $L_j = [a_j, b_j]$ };

2. For each i $(1 \le i \le k)$
   i. **Count** the number of intervals intersecting each of the intervals I_j $(1 \le j \le n)$ having nonempty intersection with L_i;
   ii. **Add** these numbers up to obtain a number r_i { r_i gives the sum of the number of intervals intersecting with intervals involved in the formation of the L_i. a is a measure of the reliability of L_i };

3. Choose the maximum of the r_i $(1 \le i \le k)$ and call it r;

4. Let $m = \min\{ i \mid r_i = r \}$ and $M = \max\{ i \mid r_i = r \}$;

5. Assign $I_p^* = [a_m, b_M]$ to be the integrated output estimate;

end.

4.3 *A Comparison of Performance*

In this new model, we find that when the number of sensors is very large, by taking the clustering of the tamely faulty sensors into consideration, we reduce the output intervals width greatly, as compared to Marzullo's [Marz 89] output interval estimate.

The Figure 3 below illustrates the superior performance of our model clearly. The numbers near each interval estimate gives the number of intervals overlapping with it. Here n = 13 and f (maximum number of fault intervals allowed) = 10. The The thick lines in Figure 3 are the intervals L_i with the numbers on them indicating their reliabilities. We may pick either the interval with higher reliability or define a range for reliabilities and pick intervals which fall in these limits. The thick line of the bottom indicates the output interval estimate for this case in Marzullo's model.

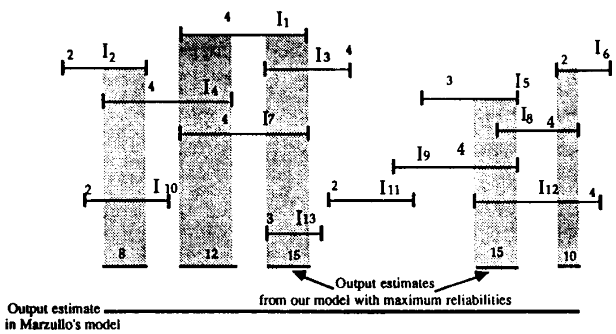

Figure 3. A compariron of performance

Table 1. Popularities of intervals

Interval	I_1	I_2	I_3	I_4	I_5	I_6	I_7	I_8	I_9	I_{10}	I_{11}	I_{12}	I_{13}
Popularity	4	2	4	4	3	2	4	4	4	2	2	4	3

In table 1, we have each interval assigned a numbering which is its popularity.

The (n - f) or more intersections (i.e. 3 to 4-intersections here) which form the output have reliabilities 8, 12, 15, 15, 10.

Thus, the advantage here is that we have the intersection weighted to help us judiciously choose the output intervals. We may employ any convenient rule depending upon our faith in the tameness of the faults to pick these intervals and enclose them by connected interval. For instance, we may choose only those intervals with maximum reliability (in this case, the intervals with reliability 15) and enclose them by a connected interval. It is clear that the worst possible width for the final output interval estimate is the smallest interval containing all intersections irrespective of their weights.

5.0 CONCLUDING REMARKS

In order to address the general problem of fault-tolerant sensor integration for a large class of sensors, it is necessary to evolve a broad-based computational framework which can accommodate a wide range of sensors and a variety of fault tolerant integration techniques depending upon the phenomenon being sensed and the method of sensing. We intend to develop a calculus of sensor integration by regarding the sensor estimates as subsets of an abstract parameter space and obtaining functional representations of the characteristics of these estimates. We then intend to obtain rules for combining these functions to get functions describing the characteristics of the output according to the kind of integration that is required to be performed.

6.0 REFERENCES

[Marz 89] Marzullo, K, "Implementing Fault-Tolerant Sensors", TR89-997, Department of Computer Science, Cornell University, May 1989.

[HaKa 90] H. Hau and R.L. Kashyap, "Belief Combination and Propagation in a Lattice Structured Inference Network," IEEE Trans. on SMC, V. 20, Jan-Feb., 1990.

[Wess 81] Wesson, R. et. al., "Network Structures for Distributed Situation Assessment", IEEE Trans. on SMC., Jan. 1981, pp. 5-23.

[BiBr 90] Blackman, S.S., and Broida, T.J., "Multiple Sensor Data Association and Fusion in Aerospace Applications", J. of Robotic systems, 7(3), 445-485 (1990).

[LuKa 89] Luo, R.C. and Kay, M.G., "Multisensor Integration and Fusion in Intelligent Systems," IEEE Trans. on SMC, V. 19, No. 5, 1989, pp. 901-931.

[HAMD 87] Henderson, T.C., Allen, P.K., Mitchie, A., Durrant-Whyte, H., Snyds, W., Eds., "Workshop on Multisensor Integration in Manufacturing Automation," Dept. of Comp. Scie., Univ. of Utah, Snowbird, Tech. Rpt. UUCS-87-006, Feb. 1987.

[LuLi 88] Luo, R.C., Lin, M. and Scherp, R.S., "Dynamic Multi-sensor Data Fusion System for Intelligent Robots," IEEE J. Robot. Automat., V. RA-4, No. 4,1988, pp. 386-396.

[Duwh 88] Durrant-Whyte, H.F., "Sensor Models and Multisensor Integration," Int. J. Robot. Res., V. 7, No. 6, 1988, pp. 97-113.