Data Incest Removal in a Survivable Estimation Fusion Architecture *

Samuel P. McLaughlin
Department of Electrical & Electronic Engineering
University of Melbourne, Aust.
s.mclaughlin@ieee.org

Robin J. Evans
Department of Electrical & Electronic Engineering
University of Melbourne, Aust.
r.evans@ee.mu.oz.au

Vikram Krishnamurthy
Department of Electrical & Computer Engineering
University of British Columbia.
vikrank@ece.ubc.ca

Abstract – An Out of Sequence Measurements (OOSM) based approach to decentralized fusion is presented with specific application to multi-sensor target tracking. A full information network architecture is considered where a fixed delay exists between nodes. The communication between nodes is a fused posterior density whose optimal conditioning set contains complete information on individual measurements from all networked sensors.

Keywords: Tracking, filtering, optimal estimation, decentralized architecture, data incest, OOSM.

1 Introduction

It is essential in any information fusion system to maintain up-to-date information. Of interest in a tracking environment is the state (position and velocity) of a specific target. Bayes rule can be applied to recursively estimate the state. These state estimates provide inputs to operational systems which respond, both offensively and defensively, to given circumstances.

A typical state estimate is in the form of a probability density. Under Gaussian noise assumptions, a density can be fully described by its mean and variance statistics. For a linear system with Gaussian noise, the Kalman filter [1] can be used to apply Bayes rule to recursively update those statistics.

Multisensor tracking and data fusion involves transmission of information across communication links. The focus of many early attempts to decentralize the Kalman filter [2], [3], [4], [5], [6], [7], [8], [9], [10] and [11] was the distribution of the computational load across multiple processors. Little attention was paid to the physical location of the sensor/fusion node nor the communication between them.

From a fusion perspective, the simplest approach to inter-node communication is to send the raw measurements. This scenario requires no processing at the source and relatively straightforward fusion techniques at the destination. However, in the presence of clutter, the physical size of this information can stretch network resources to their upper limit.

A more efficient communication would involve sending an effective measurement

\[
p(y_k | x_k) = \frac{p(x_k | Y^k)}{p(x_k | Y^{k-1})} \quad (1)
\]

which, although processed, is still unique information to a single measurement. The state estimate is the product of all effective measurements in the conditioning set.

A decentralized network employing this communication scheme is robust against individual node or communication link failure. Any information arriving at a destination node is only conditioned on information that is completely new to that node.

This approach, however, is not always possible. Legacy tracking systems exist that output a posterior density (best overall estimate) rather than processed information relating to an individual measurement. This has advantages and disadvantages. An advantage is that operational systems, that make use of the up-to-date state information, can be attached to the network without requiring a local fusion centre as an interface. A major drawback, though, is the increase
in complexity of the fusion process.

A full information network, where information is shared between all network nodes, is subject to the problem of data incest. The catalyst for this fundamentally interesting and challenging problem is the cyclic flow of repeated information.

This paper provides an introductory examination of data incest problems related to network architectures and communication topologies.

With the context of the problem introduced, Section 2 provides an overview of the data incest problem. Section 3 formulates the estimation fusion problem where a fixed delay exists between nodes. An OOSM based fusion strategy for each node is proposed in Section 4 with accompanying simulation results in Section 5. The paper is concluded in Section 6.

2 Architectures and Incest

The data incest problem stems from the multiple conditioning of an estimate on the same piece of information (measurement). The result of incest is illustrated in Figure 1. Two measurement densities are shown along with the optimal fusion of both densities. Repeat fusion of the optimal density with measurement two leads to an over confidence in a sub-optimal estimate that is biased towards the repeated information.

The decentralised network in Figure 2 is not subject to incest as to each measurement is restricted to a single path from its origin to the global fusion centre at the top of the tree. The network in Figure 3 is subject to incest and highlights two key points related to data incest:

- the need to unravel / remove information

As synchronised measurements and estimates become available they are fused immediately. Fusion of all three estimates, under an assumption of independence, results in the conditioning

\[ \mathbf{Z} = \{z_1^1, z_1^2, z_2^2, z_2^3, z_3^3\} \] (2)

due to the incestuous (multiple) conditioning on measurements from each of the three sensors. This leads to the sub-optimal posterior density

\[ p(x|\mathbf{Z}) \sim \mathcal{N}(x; \tilde{x}, \mathbf{P}). \] (3)

To obtain the optimal posterior density

\[ p(x|\mathbf{Z}) \sim \mathcal{N}(x; \hat{x}, \mathbf{P}), \] (4)

where the mean is defined as

\[ \hat{x} = E[x|\mathbf{Z}] \] (5)

with covariance

\[ \mathbf{P} = E[(x - \hat{x})(x - \hat{x})^T|\mathbf{Z}], \] (6)

the fusion process must incorporate some form of information removal to optimize the conditioning set to

\[ \mathbf{Z} = \{z_1^1, z_2^2, z_3^3\}. \] (7)
In the absence of the dotted links in Figure 3, the optimal posterior density is unobtainable from the two remaining densities as a unique solution is not available to any linear system of two equations and three unknowns. Thus, the resultant sub-optimal posterior density with conditioning set

\[ \mathbf{Z} = \{z^1, z^2, z^2, z^3\}, \]  

will be biased towards the repeated measurement \((z^2)\).

The inclusion of the third measurement fusion node, with related links, introduces a third equation without a new unknown. A unique solution is now available providing for complete incest removal with conditioning given by (7).

When any information follows multiple paths to an ultimate fusion point, (each measurement in Figure 3 follows two paths), components of that information must be unravelled at the destination fusion centre. By virtue of this, network topologies govern the ability of the system to unravel information in an optimal fashion.

The flow of information over time, in a decentralized architecture where full information is recycled between nodes, is subject to multiple paths thus making incest management an integral step in the fusion process. A fully interconnected network, with fixed delays across communications links, guarantees the ability of the system to remove incest. It still remains a non-trivial exercise to affect the removal of that incest.

Figure 4 illustrates two nodes continually integrating all available information and sending the result to the other node. The communication link between the two sensor fusion nodes is subject to a fixed time delay of one unit. At time \(k\), inputs to sensor one’s fusion centre are as follows:

- Global posterior density from node one at \(k-1\)
- Global posterior density from node two at \(k-1\)
- New measurement at node one from time \(k\)

If the fusion took place under the assumption of independence, the result at time \(k = 3\) would be sub-optimal with conditioning set

\[ \mathbf{Z} = \{z^1_1, z^2_1, z^2_1, z^1_1, z^1_2, z^2_1\}. \]  

To achieve the optimal posterior density, the redundant measurement set \(\mathbf{Z}_R = \{z^1_2, z^2_1\}\) must be removed.

When targeting a stationary object, current measurements are conditionally independent of any past measurements. This simplifies both the inclusion and removal of conditioning sets with a global set [12]. The extension to tracking a dynamic target amidst process noise, requires careful insertion of new information into the correct frame in the target’s trajectory. This leads to added complexity at the fusion centre as the presence of process noise leads to correlation measurements [13] issues among the multiple sensor nodes.

Although not considered in this paper, the data incest problem is further complicated if variable delays between nodes are introduced.

### 3 Problem Statement

Consider the standard target dynamic model

\[ \mathbf{x}_{k+1} = \mathbf{F}_k \mathbf{x}_k + \mathbf{w}_k, \quad p(\mathbf{x}_0) \sim \mathcal{N}(\mathbf{x}_0, \mathbf{P}_0) \]  

where \(\mathbf{F}\) is a known state transition matrix, \(\mathbf{x}_k \in \mathbb{R}^n\) and \(\mathbf{w}_k\) is a zero mean white noise process with known variance

\[ E[\mathbf{w}_k \mathbf{w}_j^T] = \mathbf{Q}_k \delta_{kj}. \]  

It is assumed target dynamics are static, thus yielding \(\mathbf{F}_k = \mathbf{F}\) and \(\mathbf{Q}_k = \mathbf{Q} \quad \forall \ k = \{1, \ldots, K\}\). Local measurements

\[ \mathbf{z}_n^k = \mathbf{H}_k^n \mathbf{x}_k + \mathbf{v}_n^k, \quad n = \{1, \ldots, N\} \]  

are provided by \(N\) physically remote sensors where \(\mathbf{H}_k^n\) is a known measurement matrix and \(\mathbf{v}_n^k\) is a zero mean white noise process with known variance

\[ E[\mathbf{v}_n^k \mathbf{v}_j^k] = \mathbf{R}_n^k \delta_{kj}. \]  

Measurement dynamics are also stationary providing \(\mathbf{H}_k^n = \mathbf{H}\) and \(\mathbf{R}_n^k = \mathbf{R} \quad \forall \ k = \{1, \ldots, K\}\).
The full information network shown in Figure 5 is considered, where the interconnect switch connects each node to all other nodes. The delay between nodes is fixed at a constant integer value $D$ and all measurements (illustrated by $\Delta$) are synchronized.

Given this delay, the best possible (optimal) posterior density, computed by node $n$ at time $k$, is given by
\[ p(x_k|Z^n_k) \sim \mathcal{N}(x_k; \hat{x}^n_{k[k]}, \hat{P}^n_{k[k]}) \] (14)
where the estimated mean
\[ \hat{x}^n_{k[k]} = E[x_k|Z^n_k] \] (15)
has associated error covariance
\[ \hat{P}^n_{k[k]} = E[(x_k - \hat{x}^n_{k[k]})^T(x_k - \hat{x}^n_{k[k]})|Z^n_k]. \] (16)
The conditioning set
\[ \hat{Z}^n_k = \{Z^n_{k-D}, \ldots, Z^n_{k-1}, Z^n_k, Z^n_{k+1}, \ldots, Z^n_{k+N} \} \] (17)
is a collection of complete past measurement sets
\[ Z^n_m \triangleq \{z^n_1, \ldots, z^n_m\}. \] (18)
The subscripts in (17) account for the fact that at node $n$, only local measurements are available after time $k - D$.

At each time $k$ there are are $N - 1$ arrivals
\[ A^n_m \triangleq \{p(x_k|Z^n_{k-D})\} \quad \forall \quad m = 1, \ldots, N \quad m \neq n. \] (19)
at node $n$. There is a large intersection of the conditioning sets of each of the arrivals.

The fundamental problem being presented involves updating the local posterior density with the non-redundant information from each remote arrival along with the new local measurement information.

\section{Fusion Strategy}

This section outlines a proposed strategy that fuses multiple posterior densities and measurements into a single posterior density that is free of data incest.

There are two ways to remove redundancy when fusing two estimates: before or after the combining stage. A fusion strategy that removes redundancy before combining [12] is presented for a moving target. The philosophy involves locally resolving remote measurements from remote posterior densities and storing them. At synchronised times, all locally stored information is combined to generate a global posterior density locally.

The operation of the fusion centre on board each node $n$ can be simplified to three main functional units:

- **Storage centre** - stores individual measurement information from all sensors along with the augmented prediction vector.

- **Combination centre** - combines the augmented prediction vector with a selection of measurements from all sensors to form a posterior.

- **Isolation centre** - isolates a single measurement from remotely generated posterior densities

Following an explanation of each of these units, the algorithm is presented which results in an incest free posterior density conditioned on (17).

\section{STORAGE CENTRE}

The role of the local storage centre is to ensure all information necessary for the optimal fusion of two posterior densities is available.

The key storage item
\[ Z^n_n = \begin{bmatrix} 0 & \cdots & 0 & z^n_0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & z^n_{k-D} & z^n_{k-D+1} & \cdots & z^n_{k-D} \\ \vdots & \ddots & \vdots & \ddots & \ddots & \ddots & \vdots \\ z^n_{k-2D} & \cdots & z^n_{k-2D} & z^n_{k-2D} & \cdots & z^n_{k-2D} \end{bmatrix} \] (20)
is a collection of individual measurements from all $N$ sensors. As can be seen in (20) the delay size $D$, along with $N$, dictates the required level of local storage. After updating the augmented prediction vector at time $k$, the $N$ oldest stored measurements from time $k - 2D - 1$ are discarded to make way for the single
new local measurement $z^n_k$.

The augmented prediction vector

$$\hat{X}_{k|k-2D-1} = \begin{bmatrix} \hat{x}_{i|k} \\ \vdots \\ \hat{x}_{h+1|h} \end{bmatrix},$$

(21)

with an associated error covariance

$$P_{k|k-2D-1} = \begin{bmatrix} P_{i|h} & 0 & \cdots & 0 \\ 0 & \ddots & \cdots & \vdots \\ \vdots & \cdots & 0 & P_{h+1|h} \end{bmatrix}.$$  

(22)

provides the foundation for the construction of various posterior densities. It is based upon full measurement information, from all sensors up to time $k - 2D - 1$.

Further support information is required to make each measurement workable. Matrices $R$ and $H$ serve this purpose. These constant values are known at all nodes along with $F$ and $Q$.

**COMBINATION CENTRE**

The local combination centre has three distinct purposes. It makes use of the OOSM [14] procedure of augmenting the state vector to accommodate delayed measurements in the system. The augmented prediction vector used in the combination centre

$$\hat{X}_{i|h} = \begin{bmatrix} \hat{x}_{i|h} \\ \vdots \\ \hat{x}_{h+1|h} \end{bmatrix},$$

(23)

has associated error covariance

$$P_{i|h} = \begin{bmatrix} P_{i|h} & 0 & \cdots & 0 \\ 0 & \ddots & \cdots & \vdots \\ \vdots & \cdots & 0 & P_{h+1|h} \end{bmatrix}.$$  

(24)

The augmented posterior vector is given by

$$\hat{X}_{i|j} = \begin{bmatrix} \hat{x}_{i|j} \\ \vdots \\ \hat{x}_{h+1|j} \end{bmatrix},$$

(25)

with an associated error covariance

$$P_{i|j} = \begin{bmatrix} P_{i|j} & 0 & \cdots & 0 \\ 0 & \ddots & \cdots & \vdots \\ \vdots & \cdots & 0 & P_{h+1|j} \end{bmatrix}.$$  

(26)

To convert (23) and (24) to (25) and (26) respectively requires new measurements

$$\hat{Z}_{j} = \begin{bmatrix} Z^n_j \\ \vdots \\ Z^n_{h+1} \end{bmatrix},$$

(27)

and support information

$$\hat{H}_j^n = \begin{bmatrix} H^n_j & 0 & \cdots & 0 \\ 0 & \ddots & \cdots & \vdots \\ \vdots & \cdots & 0 & H^n_{h+1} \end{bmatrix},$$

(28)

and

$$\hat{R}_j^n = \begin{bmatrix} R^n_j & 0 & \cdots & 0 \\ 0 & \ddots & \cdots & \vdots \\ \vdots & \cdots & 0 & R^n_{h+1} \end{bmatrix}.$$  

(29)

The dimension of elements $Z^n_i$, $H^n_i$ and $R^n_i \forall i \in h+1, \ldots, j$ is dependent upon the particular task being undertaken. The task also dictates the value of the time indices $h$, $i$ and $j$, although $h$ and $i$ remain fixed for set values of $n$ and $k$.

The inputs and outputs from (23) to (29) are related through

$$\hat{X}_{i|j} = \hat{X}_{i|h} + \hat{K}_j^n \hat{Y}_j^n,$$

(30)

with the covariance update taking the form

$$P_{i|j} = (I - \hat{K}_j^n \hat{H}_j^n)P_{i|h}.$$  

(31)

The innovation

$$\hat{Y}_j^n = \hat{Z}_j^n - \hat{H}_j^n \hat{X}_{i|h}$$

(32)

has covariance

$$S_j^n = \hat{H}_j^n P_{i|h}(\hat{H}_j^n)^T + \hat{R}_j^n.$$  

(33)

and Kalman gain matrix

$$\hat{K}_j^n = P_{i|h}(\hat{H}_j^n)^T (S_j^n)^{-1}.$$  

(34)

As mentioned, appropriate choice of $h$, $i$ and $j$ in (23) - (29) governs which of the three functions that the combination centre is to carry out. They include:

1. Posterior density - Output
2. Posterior density - Marginal
3. Augmented prediction vector - Update
In all three cases, the base information (augmented prediction vector) is identical with $i = k$ and $h = k - 2D - 1$. Therefore (23) and (24) are equivalent to (21) and (22) respectively, which are stored locally at all nodes.

I Posterior Density - Output

The broadcast output from node $n$ at time $k$ is the posterior density given by (14). The mean (15) and variance (16) of this posterior are the first elements of (25) and (26) respectively with $i = k$ and $j = k$.

The measurements required to achieve this, take the form of (27) with $j = k$ and

$$Z^n_l = \begin{cases} z^n_l & \text{for } k \leq l < k - D \\ \left[ \begin{array}{c} z^n_1 \\ \vdots \\ z^n_N \end{array} \right] & \text{otherwise,} \end{cases} \quad (35)$$

Similarly, the elements of the support information (28) and (29) with $j = k$ is given by

$$H^n_l = \begin{cases} H & \text{for } k \leq l < k - D \\ \left[ \begin{array}{c} H^{(1)} \\ \vdots \\ H^{(N)} \end{array} \right] & \text{otherwise} \end{cases} \quad (36)$$

and

$$R^n_l = \begin{cases} R & \text{for } k \leq l < k - D \\ \left[ \begin{array}{ccc} R^{(1)} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & R^{(N)} \end{array} \right] & \text{otherwise} \end{cases} \quad (37)$$

respectively. Note that the superscripts in brackets, included to indicate the dimension of the full matrix, are not powers.

II Posterior Density - Marginal

The marginal posterior density is used as an input to the isolation centre. The posterior density statistics of interest is

$$p(x_n|\hat{Z}^n_{k-1}) \sim N(x_n; \hat{x}^n_{k|k-1}, \hat{P}^n_{k|k-1}) \quad (38)$$

where the conditioning set

$$\hat{Z}^n_{k-1} = \{Z^n_{k-D}, \ldots, Z^n_{k-D}, Z^n_{k-1}, Z^n_{k-D}, \ldots, Z^n_{k-D} \} \quad (39)$$

are elements of (25) and (26) with $j = k - D - 1$.

This leads to use of measurements (27) with $j = k - D - 1$ and

$$Z^n_l = \begin{cases} z^n_l & \text{for } l = k - 2D \\ z^n_l & \text{otherwise,} \end{cases} \quad (40)$$

Similarly, the elements of the support information (28) and (29) with $j = k - D - 1$ is given by

$$H^n_l = \begin{cases} H^{(1)} & \text{for } l = k - 2D \\ H^{(N)} & \text{otherwise} \end{cases} \quad (41)$$

and

$$R^n_l = \begin{cases} R^{(1)} & \text{for } l = k - 2D \\ R^{(N)} & \text{otherwise} \end{cases} \quad (42)$$

respectively.

III Augmented Prediction Vector - Update

Before the oldest measurements are shifted out of (20) they must be used to update the augmented prediction vector. The outputs required are

$$\hat{X}_{k|k-2D} = \begin{bmatrix} \hat{x}_{k|k-2D} \\ \vdots \\ \hat{x}_{k-D|k-2D} \end{bmatrix}, \quad (43)$$

and

$$P^n_{i|j} = \begin{bmatrix} P^n_{i|j} & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \ldots & 0 & P^n_{i+1|j} \end{bmatrix}, \quad (44)$$

which are (25) and (26) with $j = k - D$. In this case, only $N$ measurements (27) are required

$$\hat{Z}^n_{k-2D} = Z^n_{k-2D} = \left[ \begin{array}{c} z^n_{k-2D} \\ \vdots \\ z^n_{k-2D} \end{array} \right] \quad (45)$$
with support
\[
\hat{R}^n_{k-2D} = R^N_{k-2D} = \begin{bmatrix} R^{(1)} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & R^{(N)} \end{bmatrix}
\] (47)

Using (43), the updated augmented prediction vector
\[
X_{k+1|k-2D} = \hat{F}X_{k|k-2D}
\]
\[
= \begin{bmatrix} \hat{x}_{k+1|k-2D} \\ \vdots \\ \hat{x}_{k-2D+1|k-2D} \end{bmatrix}
\] (48)
can be derived while the recursion of the covariance uses (44) to give
\[
P_{k+1|k-2D} = \hat{F}P_{k|k-2D}\hat{F}^T + Q
\]
\[
= \begin{bmatrix} P_{k+1|k-2D} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & P_{k-2D+1|k-2D} \end{bmatrix}
\] (49)

**ISOLATION CENTRE**

The purpose of the isolation centre is to extract individual remote measurements which are then placed into local storage. Using the remotely generated output and locally generated marginal posterior densities given by (14) and (38), the individual measurement can be extracted [6] by letting \( j = k - D \) to give
\[
(\hat{P}^m_{j|j})^{-1} - (\hat{P}^m_{j|j-1})^{-1} = H^TR^{-1}H
\] (50)
and
\[
\hat{x}^m_{j|j} = \hat{P}^m_{j|j} \left( (\hat{P}^m_{j|j-1})^{-1} \hat{x}^m_{j|j-1} + H^TR^{-1}z^m_j \right).
\] (51)

Rearranging (51) gives the measurement
\[
z^m_j = R(H^T)^{-1} \left( (\hat{P}^m_{j|j})^{-1} \hat{x}^m_{j|j-1} - (\hat{P}^m_{j|j-1})^{-1} \hat{x}^m_{j|j-1} \right)
\] (52)
assuming \((H^T)^{-1}\) is known and available. This remote measurement is now stored locally.

**ALGORITHM**

The algorithm that pertains to the fusion policy employed at node \( n \) at time \( k \) can be simplified as follows:

1. Update augmented prediction vector
2. Add local measurement to storage
3. For each element of arrivals:
   - generate marginal posterior density
   - isolate remote measurement
   - add remote measurement to storage
4. Create output posterior density
5. Broadcast posterior density across network

**5 Simulations**

Simulation results are currently being compiled and will be provided during the conference presentation of this paper.

**6 Conclusion**

A fusion strategy has been presented for a decentralised, full information network where a fixed delay exists between nodes. The communication between nodes is an optimal posterior density conditioned on locally available measurement information from all nodes. The inputs at each node’s fusion centre are the current local measurement, the current posterior density and incoming posterior densities from all other nodes.

**References**


