Bayesian Network Model for Data Incest in a Distributed Sensor Network *

Samuel McLaughlin
Department of Electrical and Electronic Engineering
University of Melbourne
Australia
spmcl@ee.mu.oz.au

Vikram Krishnamurthy
Department of Electrical and Computer Engineering
University of British Columbia
Canada
vikramk@ece.ubc.ca

Robin J. Evans
Department of Electrical and Electronic Engineering
University of Melbourne
Australia
r.evans@ee.mu.oz.au

Abstract – *This paper addresses the problem of linear estimation in a distributed sensor network. A fundamental issue in many-to-many sensor configurations is that of data incest, arising from the inadvertent (or ignored) multiple use of identical information. Extending recent work by the authors that considers the transmission of a global estimate, this paper casts the decentralized estimation problem in a Bayesian Network framework. A fusion strategy is presented in this framework, to address the problem of estimating a dynamic target in a network where variable delays exist between nodes.*

1 Introduction

Motivated by recent applications in network-centric (internet) systems [1] this paper extends current work by the authors in the area of data incest management within distributed sensor networks. Variable communication delays, prevalent in a packet switched network such as the internet, increase the complexity the incest management problem.

Data incest is the result of repeated use of identical information. A simple scenario, where data incest can potentially exist, is highlighted in Figure 1. Three airborne platforms (A,B,C) send track information on target (T) to a central fusion centre (F). Suppose C is out of range of F and

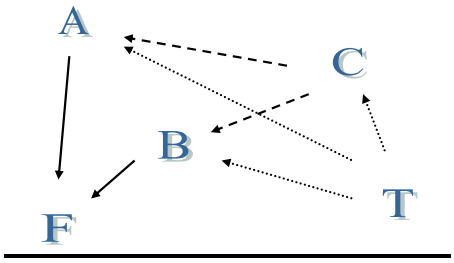


Fig. 1: POSSIBLE OCCURRENCE OF DATA INCEST

must relay information (via track-to-track fusion) through A or B without guarantee of its arrival. Imagine C can also drop in and out of range of both A and B independently. To increase the chances of information from C reaching F, C broadcasts its information to both A and B. Assuming track information from A and B is fused at F as if independent, data incest occurs when information from C reaches both A and B and is used multiple times. Table 1 outlines the three possible conditioning sets that the estimate at F might have, if the probability of transmission, between both C $\rightarrow$ A and C $\rightarrow$ B, is neither zero or one. A probability of successful transmission equal to a half, across both links, would represent a one in four chance of data incest occurring.

Table 1: Summary of possible conditioning sets at F

$C \rightarrow A$	$C \rightarrow B$	Density at F
No	No	$p(T\|A,B)$
Yes	No	$p(T\|A,B,C)$
No	Yes	$p(T\|A,B,C)$
Yes	Yes	$p(T\|A,B,C,C)$

In a fully interconnected and decentralised sensor network, where estimates based on measurements from all nodes are exchanged between nodes [2] [3], the impact of the data incest problem increases exponentially. The target in [2] was assumed stationary while the delays between nodes were variable up to a fixed maximum value. A further extension saw a non-zero target velocity considered in [3] while the communication delays were fixed at known values. This line of research continues here to provide a fusion approach applicable to tracking a moving target within a network subject to variable delays. In this paper we identify some of the issues involved, along with possible approaches to overcoming them, through casting the problem in a graphical Bayesian Network (BN) framework.

Probabilistic graphical models provide useful insight into the problem of data incest management. A Bayesian Network (BN) is a tool, in the form of a directed acyclic graph, used to model the joint distribution of a set of random variables. The graph comprises a collection of nodes, where each node represents a random variable, along with

*This work is supported by Defense Science Technology Organization (DSTO) through TDFL, the Center for Sensor Signal and Information Processing (CSSIP) and the University of Melbourne.

a set of directed edges that represent causal influence between the adjoining nodes. Readers are directed to [4], [5] for foundations on graph theoretic inference and more recently [6], [7], [8] for applications to linear dynamical systems.

Section 2 begins by formulating a particular decentralized estimation problem, in which incest management is essential in acquiring an ideal conditional estimate at each node. BNs are introduced in Section 3 as a means of representing the problem in an established graph theoretical framework. Graph manipulation techniques, that assist in overcoming the data incest problem, are also presented here. Section 4 then outlines three functional units required, within each node, to facilitate these techniques. The fusion algorithm described in Section 5 employs these units to produce an optimal global estimate. The paper concludes in Section 6.

2 Problem Formulation

Consider the standard target dynamic model

$$\mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k + \mathbf{w}_k, \qquad p(\mathbf{x}_0) \sim \mathcal{N}(\bar{\mathbf{x}}_0, \mathbf{P}_0) \tag{1}$$

where $\mathbf{F}$ is a known transition matrix for the state $\mathbf{x}_k \in \mathbb{R}^m$ and $\mathbf{w}_k$ is a zero mean white Gaussian noise process with known variance

$$E[\mathbf{w}_k \mathbf{w}_j'] = \mathbf{Q}\delta_{kj}. \tag{2}$$

Local measurements

$$\mathbf{y}_k^{(n)} = \mathbf{H}^{(n)}\mathbf{x}_k + \mathbf{v}_k^{(n)}, \qquad n = \{1, \ldots, N\} \tag{3}$$

are provided by N different sensors where $\mathbf{H}$ is the observation matrix and $\mathbf{v}_k^{(n)}$ is a zero mean white Gaussian noise process with known variance

$$E[\mathbf{v}_k^{(n)} \mathbf{v}_j^{(n)}] = \mathbf{R}^{(n)}\delta_{kj}.$$

Throughout this paper, no notational distinction is made between a random variable (RV) and its realization.

The complete set of observed measurements from all sensor nodes up to time k is defined as the *global measurement set*

$$Y^k = \{Y_k^{(n)} : n = 1, \ldots, N\} \tag{4}$$

where each subset

$$Y_k^{(n)} = \{\mathbf{y}_i^{(n)} : i = 1, \ldots, k\} \tag{5}$$

is further defined as a *node measurement set* containing all measurements from sensor node n up to time k.

Linear estimation in a centralised sensor network, where all measurements up to and including time k are available, can provide a *Globally Optimal Centralised Estimate* (GOCE)

$$\hat{\mathbf{x}}_k \triangleq E[\mathbf{x}_k | Y^k]. \tag{6}$$

In this paper, however, we consider the decentralised network shown in Figure 2. Each *sensor fusion node* (SFN) in the network consists of a local sensor coupled to a fusion centre. The interconnect switch links each SFN to *all* other SFNs in the network with a dedicated path. Information sent between any two SFNs is subject to a variable communication delay, independent of origin, destination and time sent. The following definition is used to simplify notation relating to the timing of information flows between SFNs.

Definition 2.1 *A* node-time pair, *denoted* $[n, k]$, *is a space-time reference for an event that occurs at SFN* n *at time* k.

Let $d_k^{n,m} \in 1, \ldots, \mathcal{D}$ be the communication delay for information sent from $[n, k]$, that will arrive at $[m, k + d_k^{n,m}]$. The integer $\mathcal{D}$ is defined as the maximum delay between any two SFNs in the system. An important assumption we make for this paper is that there is no overlap in the communication trajectories either directly or indirectly.

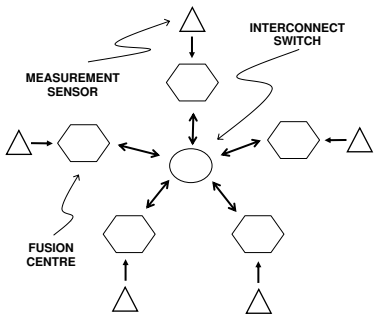


Fig. 2: FULL INFORMATION NETWORK

Let $Z_k^{(m)} = Y_i^{(m)}$ be the set of all *available* measurements at SFN n originating from SFN m up to some time $i < k$. Denoting

$$\mathcal{Z}_k^{(n)} = \{Z_k^{(m)} : m = 1, \ldots, N\} \tag{7}$$

as the *available* measurement set at $[n, k]$ implies $\mathcal{Z}_k^{(n)} \subseteq Y^k$. Information must have *arrived* from SFN m, for $m \in 1, \ldots, N$ and $m \neq n$, to be available at SFN n. We assume information sent from $[m, j]$, for $j < k$, can arrive at SFN n in the continuous interval $k - 1$ to k. We only consider estimation updates at discrete intervals and thus formalise the concept of an arrival to reflect this.

Definition 2.2 *An input element is said to* arrive *at the fusion centre associated with* $[n, k]$ *if it is available as an input at* $[n, k]$ *but it is not available at* $[n, k - 1]$.

Given this arrival requirement, the GOCE is unlikely to be available at any decentralised nodes. Thus, the optimal posterior density computed at $[n, k]$, is given by

$$p(\mathbf{x}_k | \mathcal{Z}_k^n) \quad \sim \quad \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}^n, \mathbf{P}_{k|k}^n), \tag{8}$$

where the *Globally Optimal Decentralized Estimate* (GODE)

$$\hat{\mathbf{x}}^n_{k|k} \triangleq E[\mathbf{x}_k|\mathcal{Z}^n_k] \tag{9}$$

has associated error covariance

$$\mathbf{P}^n_{k|k} \triangleq E\Big[(\mathbf{x}_k - \hat{\mathbf{x}}^n_{k|k})(\mathbf{x}_k - \hat{\mathbf{x}}^n_{k|k})^T|\mathcal{Z}^n_k\Big]. \tag{10}$$

Ideally, this statistical pair alone is to be sent to all other SFNs in the network. We find, as will be shown later, that an additional vector must be attached to each pair before being sent.

The fusion process at $[n,k]$ has two forms of statistical input. The first is the new local measurement, given by (3), which is preprocessed to provide a likelihood density. The second is a set, $p^{(n)}_{IN}(k)$, which consists of pairs of Gaussian statistics that represent posterior densities of the form given by (8). Given the variable delay between nodes, $p^{(n)}_{IN}(k)$ may contain a maximum $(N-1)\times\mathcal{D}$ or minimum zero elements.

3 Estimation Using Bayesian Networks

We begin this section with a brief introduction to the structure of Bayesian Networks (BNs). We then apply BN concepts to centralised estimation, as a precursor to investigating the data incest problem encountered in decentralised systems. The section concludes with an outline of two operations that can be used to manipulate a pair of BNs.

3.1 Belief Propagation in Bayesian Networks

Let $\mathcal{G}$ be a directed acyclic graph [9] where each node in the set V represents a multivariate random variable from the set $\mathbf{u} = \{\mathbf{u}_1, \ldots, \mathbf{u}_W\}$. Where a link exists between two nodes, it is represented by a conditional distribution of $\mathbf{u}_v$ given its parents $\mathbf{u}_{pa(v)}$ for each $v \in V$. Setting the density to $p(\mathbf{u}_v|\mathbf{u}_{pa(v)})$, the causal Markov property prevalent on a BN factorises the joint density on $\mathcal{G}$ to

$$\begin{aligned} p(\mathbf{u}) &= p(\mathbf{u}_1, \ldots, \mathbf{u}_W) \\ &= \prod_{v\in V} p(\mathbf{u}_v|\mathbf{u}_{pa(v)}), \end{aligned} \tag{11}$$

thereby greatly simplifying the calculation of $p(\mathbf{u})$.

State estimation is generally concerned with finding the conditional density of an individual RV, based on a set of observed variables (evidence or measurements) which form a subset of $\mathbf{u}$. Inference on a BN provides this conditional density and the belief propagation algorithm [4] provides an efficient mechanism for the distributing evidence through the BN. Top-down propagation follows in the direction of the arrows while bottom-up propagation involves information flow which opposes the direction of the arrows.

Let Y denote the set of all observed information on a graph $\mathcal{G}$ containing the (non leaf) node X. Let Y^-_X represent evidence found below node X (following a positive arrow direction from X) and let Y^+_X be the remaining observed information on $\mathcal{G}$ (following a negative arrow direction from X). From [4] we know

$$\begin{aligned} p(X|Y) &= \alpha p(X|Y^+_X)p(Y^-_X|X) \\ &= \alpha\pi(X)\lambda(X) \end{aligned} \tag{12}$$

where α is a normalising constant that has no impact on either of the two Gaussian statistics. The quantities $\pi(X)$ and $\lambda(X)$ represent messages passed through the network as posterior information (top-down) and likelihood information (bottom-up) respectively.

Consider a node X with children $C_1, \ldots, C_n$. The collective likelihood is

$$\lambda(X) = \prod_k \lambda_{C_k}(X). \tag{13}$$

The contribution provided by each individual child likelihood is

$$\lambda_{C_k}(X) = \sum_{C_k} \lambda(C_k)P(C_k|X). \tag{14}$$

The calculation of $\pi(X)$ considers only information passed from each parent D, giving the prediction

$$\pi(X) = \sum_D P(X|D)\pi_X(D) \tag{15}$$

where

$$\pi_X(D) = \alpha\pi(D)\prod_{S_X} \lambda_{S_X}(D) \tag{16}$$

is the posterior density passed from D and S_X denotes a node that shares the same parent as X. Figure 3 outlines the four possible message passing options, two up and two down, that are applicable to the node X.

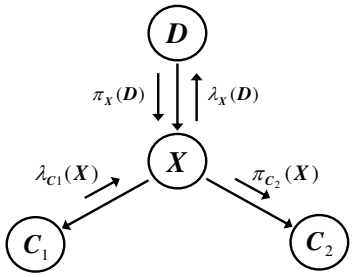


Fig. 3: MESSAGE PASSING OPTIONS AT NODE X

We now consider applying this message passing philosophy to the problem of centralised and decentralised estimation.

3.2 Centralised Estimation

In a prediction and filtering context, where smoothing is not considered and the desired output is given by (6), only

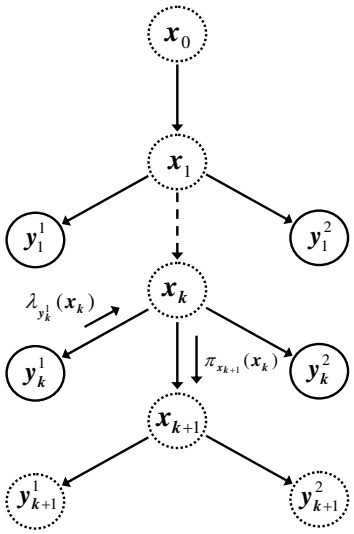


Fig. 4: CENTRALIZED BAYESIAN NETWORK $\mathcal{G}_k$

two of four possible message passing directions a BN $\mathcal{G}$ are required. Figure 4 illustrates a tree structure that models equations (1) and (3) for $N = \mathcal{D} = 2$. Unobserved nodes are represented by dotted circles. Bottom-up propagation of $\lambda_{\mathbf{x}_k}(\mathbf{x}_{k-1})$ from $\mathbf{x}_k$ to $\mathbf{x}_{k-1}$ (smoothing) is not considered nor is $\pi_{\mathbf{y}_k^{(n)}}(\mathbf{x}_k)$ as no inference is performed on unobserved leaf nodes $\mathbf{y}_k^{(n)}$ through top-down propagation from $\mathbf{x}_k$.

The process of inferring a conditional estimate of $\mathbf{x}_k$ begins with the initialisation of the root node

$$\pi(\mathbf{x}_0) = \left(\frac{1}{\sqrt{2\pi}}\right)^r |\mathbf{R}|^{-1/2} \times exp\left[-\frac{1}{2}(\mathbf{x}_0 - \mu_0)'\mathbf{R}^{-1}(\mathbf{x}_0 - \mu_0)\right]. \quad (17)$$

All likelihood densities $\lambda_{\mathbf{y}_k^{(m)}}(\mathbf{x}_k)$ relating to measurement nodes are set to 1 reflecting their initial unobserved state. Computationally, this is achieved by assigning each λ a covariance matrix with infinite values on the diagonal and 0 elsewhere and equating the mean to $\mathbf{0}$.

As the estimation process moves through time to k, the traditional steps of prediction and updating [10] are modelled by top-down propagation ($\mathbf{x}_{k-1} \rightarrow \mathbf{x}_k$) and bottom-up propagation ($\mathbf{y}_k^{(n)} \rightarrow \mathbf{x}_k$) respectively [6].

The prediction step involves propagating the posterior message $\pi_{\mathbf{x}_k}(\mathbf{x}_{k-1}) = N_r(\mathbf{x}_{k-1}; \mu_\pi^-, \Sigma_\pi^-)$ down the link $\mathbf{x}_{k-1} \rightarrow \mathbf{x}_k$, which is governed by the system equation

$$p(\mathbf{x}_k|\mathbf{x}_{k-1}) = \left(\frac{1}{\sqrt{2\pi}}\right)^r |\mathbf{R}|^{-1/2} \times exp\left[-\frac{1}{2}(\mathbf{x}_k - \mathbf{F}\mathbf{x}_{k-1})'\mathbf{R}^{-1}(\mathbf{x}_k - \mathbf{F}\mathbf{x}_{k-1})\right]. \quad (18)$$

Using (15) with a single parent $\mathbf{x}_k$, gives the predicted density

$$\begin{aligned}\pi(\mathbf{x}_k) &= \int_{\mathbf{x}_{k-1}} p(\mathbf{x}_k|\mathbf{x}_{k-1})\pi_{\mathbf{x}_k}(\mathbf{x}_{k-1})d\mathbf{x}_{k-1} \\ &= N_r(\mathbf{x}_k; \mu_\pi, \Sigma_\pi)\end{aligned} \quad (19)$$

where

$$\Sigma_\pi = \mathbf{Q} + \mathbf{F}\Sigma_\pi^-\mathbf{F}' \quad (20)$$

and

$$\mu_\pi = \mathbf{F}\mu_\pi^- \quad (21)$$

The update stage involves a bottom-up propagation opposing the link $\mathbf{x}_k \rightarrow \mathbf{y}_k^{(n)}$, which is defined by the measurement equation

$$p(\mathbf{y}_k^{(n)}|\mathbf{x}_k) = \left(\frac{1}{\sqrt{2\pi}}\right)^r |\mathbf{R}|^{-1/2} \times exp\left[-\frac{1}{2}(\mathbf{y}_k^{(n)} - \mathbf{H}^{(n)}\mathbf{x}_k)'\mathbf{R}^{-1}(\mathbf{y}_k^{(n)} - \mathbf{H}^{(n)}\mathbf{x}_k)\right] \quad (22)$$

Using (13), the collective likelihood of the two new measurements (which become observed on $\mathcal{G}$) at time k is

$$\begin{aligned}\lambda(\mathbf{x}_k) &= \lambda_{\mathbf{y}_k^{(1)}}(\mathbf{x}_k)\lambda_{\mathbf{y}_k^{(2)}}(\mathbf{x}_k) \\ &= N_r(\mathbf{x}_k; \mu_\lambda, \sigma_\lambda)\end{aligned} \quad (23)$$

where

$$\Sigma_\lambda = [\mathbf{H}^{(n)'}\mathbf{R}^{(n)-1}\mathbf{H}^{(n)} + \mathbf{H}(n)'\mathbf{R}^{(n)-1}\mathbf{H}^{(n)}]^{-1} \quad (24)$$

and

$$\mu_\lambda = \Sigma_\lambda[\mathbf{H}^{(n)}\mathbf{R}^{(n)-1}\mathbf{y}_k^{(1)} + \mathbf{H}^{(n)}\mathbf{R}^{(n)-1}\mathbf{y}_k^{(2)}]. \quad (25)$$

Combining (19) and (23) through the use of (12) gives the updated posterior density

$$\begin{aligned}p(\mathbf{x}_k|Y^k) &= \pi_{\mathbf{x}_{k+1}}(\mathbf{x}_k) \\ &= \alpha\lambda(\mathbf{x}_k)\pi(\mathbf{x}_k) \\ &= N_r(\mathbf{x}_k; \mu_{\mathbf{x}_k}, \Sigma_{\mathbf{x}_k})\end{aligned} \quad (26)$$

where

$$\Sigma_{\mathbf{x}_k} = [\Sigma_\lambda^{-1} + \Sigma_\pi^{-1}]^{-1} \quad (27)$$

and

$$\mu_{\mathbf{x}_k} = \Sigma_{\mathbf{x}_k}[\Sigma_\lambda^{-1}\mu_\lambda + \Sigma_\pi^{-1}\mu_\pi] \quad (28)$$

become inputs to the next prediction step.

3.3 Decentralised Estimation

The fundamental difference between decentralized and centralized estimation is the computation of a global estimate locally at each SFN. Accommodating a unique GODE at each SFN n, requires the local administration of a unique BN $\mathcal{G}_k^{(n)}$. Recall that the GODE and its covariance, given by (9) and (10) respectively, are transmitted between SFNs. We define this pairing of parameters as the *output message* (OM) for $[m, j]$, which we denote

$$\pi_{\mathbf{x}_{j+1}}^{(n)}(\mathbf{x}_j) = p(\mathbf{x}_j|\mathcal{Z}_j^{(m)}). \quad (29)$$

Delays in the transmission of the OM between SFNs will result in unobserved leaf nodes of the destination BN $\mathcal{G}_k^{(n)}$. However, the subgraph on $\mathcal{G}_k^{(n)}$, between $\mathbf{x}_{k-D-1}$ and $\mathbf{x}_0$, will resemble the centralised graph shown in Figure 4. These two scenarios are shown in Figure 5 for the case where $N = \mathcal{D} = 2$

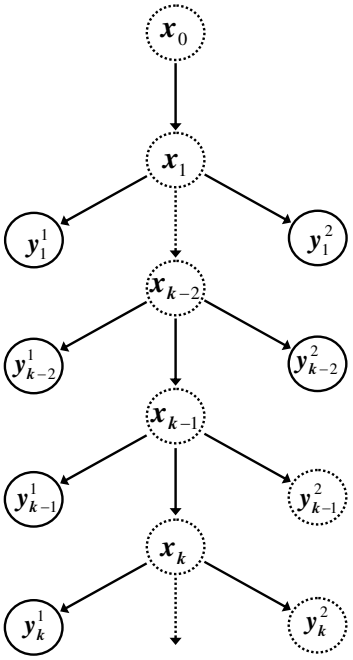


Fig. 5: DECENTRALIZED BAYESIAN NETWORK $\mathcal{G}_k^{(n)}$

Decentralised estimation at $[n, k]$ requires fusing multiple BNs and inferring the next OM on the resultant $\mathcal{G}_k^{(n)}$. The cyclic nature of the data flows demands caution be exercised during the fusion process, so that data incest is avoided. The presence of data incest in a BN framework, is reflected in Figure 6, by the repeat of the observed node $\mathbf{y}_k^{(1)}$.

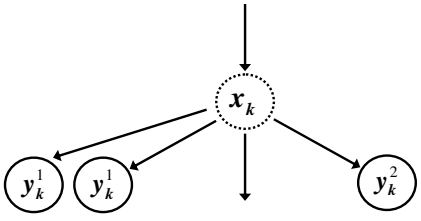


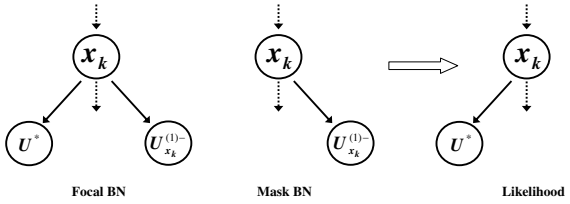
Fig. 6: INCESTED BAYESIAN NETWORK

The remainder of this section outlines two fundamental operations on BNs. These operations can be used to combat the problem of incest and facilitate the optimal fusion of multiple BNs.

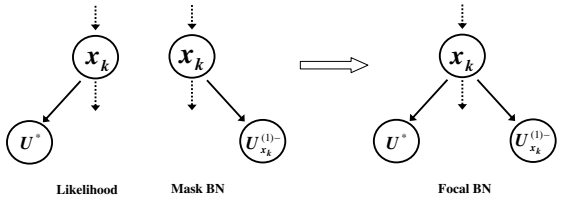
3.4 BN Filter Operations

The optimal fusion of two BNs, which contain information common to both, requires two separate operations. Figure 3.4 illustrates the general principle behind both the *isolation* and *combination* procedure on two BNs.

Let $\mathcal{G}_1$ and $\mathcal{G}_2$ be two arbitrary BNs with node sets related through $V_1 \subset V_2 = \{V_1, V^*\}$, where V^* represents a single leaf node observed in $\mathcal{G}_2$ but not in $\mathcal{G}_1$. $\mathcal{G}_2$ is defined as the *focal* BN and $\mathcal{G}_1$ is the *mask* BN of $\mathcal{G}_2$. Isolation involves the use of the mask to extract unique information, relating to V^*, from the focal BN. Two messages



(a) Isolation of Bayesian Networks



(b) Combination of Bayesian Networks

Fig. 7: BAYESIAN NETWORK OPERATIONS

are considered known on each BN. The first is the predicted density $\pi^{(i)}(\mathbf{x}_k)$ which is identical for both BNs. The second are the posterior densities $\pi^{(i)}_{\mathbf{x}_{k+1}}(\mathbf{x}_k) = p(\mathbf{x}_k|U_k^{(i)})$ for each BN where $U_k^{(i)}$ represents the set of observed leaf nodes on $\mathcal{G}_i$ for $i = 1, 2$. The set $U_k^{(i)}$ contains evidence from above and below $\mathbf{x}_k$ represented via (12) as $U_k^{(i)} = \{U^{(i)+}_{\mathbf{x}_k}, U^{(i)-}_{\mathbf{x}_k}\}$. From the deifinition of the relative node sets, it can be shown that $U^{(1)+}_{\mathbf{x}_k} = U^{(2)+}_{\mathbf{x}_k}$ and $U^{(2)-}_{\mathbf{x}_k} = \{U^{(1)-}_{\mathbf{x}_k}, U^*\}$ where U^* is a RV associated with node V^*. Applying (12) gives

$$\begin{aligned} p(\mathbf{x}_k|U_k^{(1)}) &= \alpha p(\mathbf{x}_k|U^{(1)+}_{\mathbf{x}_k}) p(U^{(1)-}_{\mathbf{x}_k}|\mathbf{x}_k) \\ &= \alpha \pi^{(1)}(\mathbf{x}_k)\lambda^{(1)}(\mathbf{x}_k) \end{aligned} \tag{30}$$

and

$$\begin{aligned} p(\mathbf{x}_k|U_k^{(2)}) &= \alpha p(\mathbf{x}_k|U^{(2)+}_{\mathbf{x}}) p(U^{(2)-}_{\mathbf{x}}|\mathbf{x}_k) \\ &= \alpha \pi^{(2)}(\mathbf{x}_k)\lambda^{(2)}(\mathbf{x}_k) \\ &= \alpha \pi^{(1)}(\mathbf{x}_k)\lambda^{(1)}(\mathbf{x}_k)\lambda_{U^*}(\mathbf{x}_k) \end{aligned} \tag{31}$$

Division of (31) by (30)

$$\begin{aligned}\frac{p(\mathbf{x}_k|U_k^{(2)})}{p(\mathbf{x}_k|U_k^{(1)})} &= \frac{\alpha p(\mathbf{x}_k|U_{\mathbf{x}_k}^{(2)+})p(U_{\mathbf{x}_k}^{(2)-}|\mathbf{x}_k)}{\alpha p(\mathbf{x}_k|U_{\mathbf{x}_k}^{(1)+})p(U_{\mathbf{x}_k}^{(1)-}|\mathbf{x}_k)} \\ &= \frac{\alpha\pi^{(1)}(\mathbf{x}_k)\lambda^{(1)}(\mathbf{x}_k)\lambda_{U^*}(\mathbf{x}_k)}{\alpha\pi^{(1)}(\mathbf{x}_k)\lambda^{(1)}(\mathbf{x}_k)} \\ &= \lambda_{U^*}(\mathbf{x}_k) \\ &= p(U^*|\mathbf{x}_k)\end{aligned} \tag{32}$$

isolates the likelihood for the measurement associated with U^*.

Combination is simply the reverse of the isolation operation and is equivalent to observing evidence related to U^*, on the graph $\mathcal{G}_2$. Combining $\lambda_{U^*}(\mathbf{x}_k)$ with $\pi^{(1)}_{\mathbf{x}_{k+1}}(\mathbf{x}_k)$ gives

$$\begin{aligned}\lambda_{U^*}(\mathbf{x}_k)\pi^{(1)}_{\mathbf{x}_{k+1}}(\mathbf{x}_k) &= \alpha\lambda_{U^*}(x_k)\pi^{(1)}(\mathbf{x}_k)\lambda^{(1)}(\mathbf{x}_k) \\ &= \alpha\pi^{(2)}(\mathbf{x}_k)\lambda^{(2)}(\mathbf{x}_k) \\ &= \alpha p(\mathbf{x}_k|U_{\mathbf{x}_k}^{(2)+})p(U_{\mathbf{x}_k}^{(2)+}|\mathbf{x}_k) \\ &= p(\mathbf{x}_k|U_k^{(2)})\end{aligned} \tag{33}$$

4 Fusion Centre Components

The following section describes the components that make up each fusion centre. The primary aim at $[n,k]$ is to infer the OM given by (29) on $\mathcal{G}_k^{(n)}$. To meet this aim, the fusion centre consists of three main functional units: a storage centre, an isolation centre and a combination centre. For readability purposes, where a superscript is omitted, the quantity relates to the local SFN n under consideration.

4.1 Storage Centre

In a BN framework, a decentralised estimation problem requires separate BNs to be administered at each SFN. When filtering, where only a subset of all message passing options are used, full information (the complete set of messages on the BN) need not be stored. The minimum requirements for local storage are governed by the maximum delay $\mathcal{D}$, the number N of SFNs in the system and the dimension of the state vector r.

The largest local storage element is the *recent likelihood message* (RLM) set, denoted Λ_k, which contains all *observed* measurement likelihood messages $\lambda_{\mathbf{y}_j^{(m)}}(\mathbf{x}_j)$ on $\mathcal{G}_k$ for $m \in 1,\ldots,N$ and $j \in k-2\mathcal{D},\ldots,k$. Let

$$\lambda_j = \{m : \lambda_{\mathbf{y}_j^{(m)}}(\mathbf{x}_j) \in \Lambda_k\} \tag{34}$$

be the index set of observed measurements that are children of $\mathbf{x}_j$. The RLM can then be written as

$$\Lambda_k = \{\lambda_{\mathbf{y}_j^{(m)}}(\mathbf{x}_j) : m \in \lambda_j, j = k-2\mathcal{D},\ldots,k\}. \tag{35}$$

Elements of Λ_k are used in all three functions mentioned above.

The remaining information storage is provided to limit the computation time for each OM. In the process, the size of the RLM is also limited. The set of *delayed posterior messages* (DPM) is denoted

$$\Pi_k = \{\bar{\pi}_{\mathbf{x}_i}(\mathbf{x}_{i-1}) : i = k-\mathcal{D}, k-2\mathcal{D}\} \tag{36}$$

where each element $\bar{\pi}_{\mathbf{x}_i}(\mathbf{x}_{i-1})$ is conditioned on the global measurement set Y^{i-1}.

4.2 Isolation Centre

This function block is specifically used to extract non-redundant information, present in the OMs that arrive from other SFNs, in order to update the RLM. For each incoming OM $\pi^{(m)}_{j+1}(\mathbf{x}_j)$, with conditioning set $\mathcal{Z}_j^{(m)}$, the combination centre (as will be shown) provides a *masking agent* (MA) denoted $\tilde{\pi}^{(m)}_{\mathbf{x}_{j+1}}(\mathbf{x}_j)$. The MA conditioning set $\tilde{\mathcal{Z}}_j^{(m)}$, is such that $\mathcal{Z}_j^{(m)} = \{\tilde{\mathcal{Z}}_j^{(m)}, \mathbf{y}_j^{(m)}\}$. The set $\tilde{\Lambda}_k$ is said to contain likelihood messages corresponding to elements of $\tilde{\mathcal{Z}}_j^{(m)}$. By definition

$$\begin{aligned}\pi^{(m)}_{\mathbf{x}_{j+1}}(\mathbf{x}_j) &= \alpha p(\mathbf{x}_j|\mathcal{Z}_j^{(m)+})p(\mathcal{Z}_j^{(m)-}|\mathbf{x}_j) \\ &= \alpha\pi^{(m)}(\mathbf{x}_j)\lambda^{(m)}(\mathbf{x}_j)\end{aligned} \tag{37}$$

and

$$\begin{aligned}\tilde{\pi}^{(m)}_{\mathbf{x}_{j+1}}(\mathbf{x}_j) &= \alpha p(\mathbf{x}_j|\tilde{\mathcal{Z}}_j^{(m)+})p(\tilde{\mathcal{Z}}_j^{(m)-}|\mathbf{x}_j) \\ &= \alpha\tilde{\pi}^{(m)}(\mathbf{x}_j)\tilde{\lambda}^{(m)}(\mathbf{x}_j).\end{aligned} \tag{38}$$

Given the conditioning sets and (13), we find that $\pi(\mathbf{x}_j) = \tilde{\pi}(\mathbf{x}_j)$ and $\lambda(\mathbf{x}_j) = \tilde{\lambda}(\mathbf{x}_j)\lambda_{\mathbf{y}_j^{(m)}}(\mathbf{x}_j)$. From (32), division of (37) by (38) will extract the likelihood message

$$\lambda_{\mathbf{y}_j^{(m)}}(\mathbf{x}_j) = \frac{\pi^{(m)}_{\mathbf{x}_{j+1}}(\mathbf{x}_j)}{\tilde{\pi}^{(m)}_{\mathbf{x}_{j+1}}(\mathbf{x}_j)} \tag{39}$$

4.3 Combination Centre

The combination centre at $[n,k]$ is responsible for four functional outputs. The primary output is obviously the OM. A secondary output is the MA, which is used as an input to the isolation centre. The remaining two outputs, which are the elements of the DPM, both result from interim messages in the provision of the primary and secondary outputs.

Calculation of OM

$\mathcal{D}+1$ stages of message passing are required to infer the OM on $\mathcal{G}_k$. Each stage comprises a top-down and bottom-up message being passed to the focal node $\mathbf{x}_j$ for that stage. The top-down message $\pi_{\mathbf{x}_j}(\mathbf{x}_{j-1})$ is conveyed to $\mathbf{x}_j$ via (19), giving the predicted message $\pi(\mathbf{x}_j)$. Using (13) and a subset of elements from Λ_k, the bottom-up message

$$\lambda(\mathbf{x}_j) = \prod_{i\in\lambda_j} \lambda_{\mathbf{y}_j^{(i)}}(\mathbf{x}_j) \tag{40}$$

contributes information from all observed measurements, having originated at time j, that are available at $[n,k]$. Using (26), $\lambda(\mathbf{x}_j)$ and $\pi(\mathbf{x}_j)$ combine to give the stage output message $\pi_{x_{j+1}}(\mathbf{x}_j)$.

Upon completion of $\mathcal{D}+1$ of these message passing stage, the final message provides the desired OM $\pi_{\mathbf{x}_{k+1}}(\mathbf{x}_k)$.

Calculation of MA

The approach applied to a MA is similar to that employed for the OM. The key differences are in the number of message passing stages and the set of observed measurements used at each of those stages.

Let $\tilde{\pi}^{(m)}_{\mathbf{x}_{j+1}}(\mathbf{x}_j)$, for $k-\mathcal{D}-1<j<k$, be the output MA to be computed at $[n,k]$. The number of message passing stages required to reach the MA is $\tilde{\mathcal{D}}=j-(k-2\mathcal{D})+1$. Let $\mathbf{x}_i$ be the focal node at each of these stages. The prediction at each stage is the same as for the OM, with (19) mapping $\tilde{\pi}^{(m)}_{\mathbf{x}_i}(\mathbf{x}_{i-1})\mapsto\tilde{\pi}^{(m)}(\mathbf{x}_i)$. For $i=j-2\mathcal{D},\ldots,j$, the creation of the bottom up message $\tilde{\lambda}(\mathbf{x}_i)$ necessarily differs to that of the OM as the MA contains only a subset of the observed measurements available at $[n,k]$. By defining

$$\tilde{\lambda}_i=\{o:\lambda_{\mathbf{y}_i^{(o)}}(\mathbf{x}_i)\in\tilde{\Lambda}_k\},\tag{41}$$

we can write

$$\tilde{\lambda}(\mathbf{x}_i)=\prod_{a\in\tilde{\lambda}_i}\lambda_{\mathbf{y}_i^{(a)}}(\mathbf{x}_i).\tag{42}$$

However, a means by which to check the condition in (41) must be introduced.

Definition 4.1 *The* message composition vector *(MCV) for OM $\pi^{(n)}_{\mathbf{x}_{k+1}}(\mathbf{x}_k)$, denoted*

$$\Gamma^{(n)}_k=\{\gamma^{(l)}_k:l=1,\ldots,N\}\tag{43}$$

is an N dimensional vector that indicates the composition of the conditioning set $\mathcal{Z}^{(n)}_k$. Each integer element $\gamma^{(l)}_k$ represents the time of the most recent likelihood, available at $[n,k]$, that originated from SFN l.

The assumption preventing any communication overlap enforces the condition $\gamma^{(n)}_k\geqslant\gamma^{(n)}_h$ for $k>h$.

Using elements of $\Gamma^{(m)}_j$ for the MA, the conditioning set can be written as

$$\mathcal{Z}^{(m)}_j=\{Y^{(l)}_{\gamma^{(l)}_k}:l=1,\ldots,N\},\tag{44}$$

where $Y^{(l)}_{\gamma^{(l)}_k}$ is the node measurement set given by (5). Establishing the elements of $\tilde{\lambda}_j$, under the condition of (41), also makes use of the related MCV. For $l\neq m$, $\lambda^{(l)}_{\mathbf{y}_i}(\mathbf{x}_i)\in\tilde{\lambda}_i$ iff $\gamma^{(l)}_i\geqslant i$. For $l=m$, $\lambda^{(l)}_{\mathbf{y}_i}(\mathbf{x}_i)\in\tilde{\lambda}_i$ iff $\gamma^{(l)}_i>i$. These two conditions ensure that $\mathcal{Z}^{(m)}_j=\{\tilde{\mathcal{Z}}^{(m)}_j,\mathbf{y}^{(m)}_j\}$.

Upon conclusion of $\tilde{\mathcal{D}}>\mathcal{D}$ message passing stages, the final message provides the desired MA $\tilde{\pi}^{(m)}_{\mathbf{x}_{j+1}}(\mathbf{x}_j)$.

Updating the DPM

The first element of the set Π_{k+1} corresponds to a stage output message in the combination process for the OM. Let $d=k-\mathcal{D}$. The delay characteristics guarantee that λ_d (relevant during the first stage of the OM calculation) contains the maximum N indices. This provides

$$\bar{\pi}_{\mathbf{x}_d}(\mathbf{x}_{d-1})=\pi_{\mathbf{x}_d}(\mathbf{x}_{d-1}).\tag{45}$$

The second element of the set Π_{k+1} corresponds to a stage output message in the calculation of a MA. Let $e=k-2\mathcal{D}$. Consider the calculation of $\tilde{\pi}_{\mathbf{x}_j}(\mathbf{x}_{j-1})$ for any $j\geqslant k-\mathcal{D}$. Again, the delay characteristics guarantee the set $\tilde{\lambda}_e$ (relevant to the first stage of the MA calculation) contains the maximum N indices. This provides

$$\bar{\pi}_{\mathbf{x}_e}(\mathbf{x}_{e-1})=\tilde{\pi}^{(m)}_{\mathbf{x}_e}(\mathbf{x}_{e-1}).\tag{46}$$

5 Fusion Algorithm

This section provides a brief outline of a possible fusion approach that ensures an optimal (incest free) global estimate is computed at $[n,k]$. The algorithm integrates the functional units just described to provide $\pi_{\mathbf{x}_{k+1}}(\mathbf{x}_k)$. Important considerations relating to the implementation of the algorithm are also identified.

Before looking closely at the $\mathcal{D}$ stages required to calculate the OM, consideration must be given to processing the input elements of $p^{(n)}_{IN}(k)$. For each incoming OM $\pi^{(m)}_{\mathbf{x}_{j+1}}(\mathbf{x}_j)$, the combination centre provides the MA $\tilde{\pi}^{(m)}_{\mathbf{x}_{j+1}}(\mathbf{x}_j)$. In providing the MA, the first top-down message is always provided by the second element of Π_k. Both the MA and the incoming OM are fed into the isolation centre, which outputs the likelihood $\lambda_{\mathbf{y}^{(m)}_j}(\mathbf{x}_j)$. It is necessary to process the input OMs in chronological order beginning with the eldest, as later MAs may rely on earlier isolations having already occurred. The sequence of incoming OMs can be processed as a batch, but computation time for the OM at $[n,k]$ is reduced if the isolation is effected in equivalent stages to the OM.

Each stage of the OM calculation begins with the standard prediction stage mapping $\pi_{\mathbf{x}_j}(\mathbf{x}_{j-1})\mapsto\pi(\mathbf{x}_j)$ using (19). Aligning isolation blocks with OM stages implies that within each OM stage, there $a\geqslant 0$ MA outputs generated and isolations performed. The number a corresponds to the number of arrivals $\pi^{(m)}_{\mathbf{x}_{j+1}}(\mathbf{x}_j)\in p^{(n)}_{IN}(k)$ that contain unique information relating to the focal stage node $\mathbf{x}_j$.

The likelihoods resulting from the isolation are placed in Λ_k and $\Gamma^{(n)}_k$ must then be updated accordingly. Once all inputs originating from time j are isolated and stored, all

applicable likelihoods for that OM stage can be collected and combined using (40). The result is then fused with $\pi(\mathbf{x}_j)$ to give the stage output message $\pi_{\mathbf{x}_{j+1}}(\mathbf{x}_j)$.

The first OM stage involves the focal node $\mathbf{x}_j$ for $j = k - \mathcal{D}$. $\Gamma_k^{(n)}$ is inherited from $[n, k-1]$ and (like Λ_k) will be updated, after each isolation, throughout the course of the $\mathcal{D}$ stages. The first top down message $\pi_{\mathbf{x}_j}(\mathbf{x}_{j-1})$ is provided by the first element of Π_k.

Following the top-down message pass, any isolations required on inputs must be carried out. Once this is complete, the nature of the delays guarantee that λ_j contains all N index values. As a result, the output message from the first OM stage not only provides the top-down message for stage two but also constitutes the first element of Π_{k+1}.

The second element of Π_{k+1} will also be updated during this first OM stage if in fact any inputs originating from j were isolated during the stage. If not, the first OM stage to perform an isolation will provide the update. This time, the delay characteristics ensure that $\tilde{\lambda}_{k-2\mathcal{D}}$ contains all N indices. Thus, the output message from the first MA stage provides the second element of Π_{k+1}.

The minimum delay of 1 dictates that the final stage of any MA process will not have any observed nodes. Consequently there is no bottom-up propagation. The final stage of the OM process contains one observed node, representing the new local measurement $\mathbf{y}_k^{(n)}$. Once this measurement is incorporated following prediction, through $\lambda_{\mathbf{y}_k^{(n)}}(\mathbf{x}_k)$ the final message (or OM) $\pi_{\mathbf{x}_{k+1}}(\mathbf{x}_k)$ is then sent to all other nodes.

6 Conclusion

The aim of this paper was to present an overview of the problem of data incest in decentralized data incest problems. The particular situation considered involved the many-to-many transmission of optimal global estimates. The estimation problem was cast in a BN framework to provide a simplified description of the operation of the fusion centre at each SFN. It remains to be seen whether the increase in accuracy warrants the increase in processing time, although it should be pointed out that this solution is applicable to a large range of linear estimation problems.

References

[1] L. J Sciacca and R. J. Evans. Cooperative sensor networks with bandwidth constraints. In *Proceedings of Aerosense SPIE Conference, Battlespace Digitisation and Network-Centric Warfare*, Orlando Florida, April 2002.

[2] S. McLaughlin, V. Krishnamurthy, and S. Challa. Managing data incest in a distributed sensor network. In *Proceedings of ICASSP*, Hong Kong, April 2003.

[3] S. McLaughlin, R. Evans, and V. Krishnamurthy. Data incest removal in a survivable estimation fusion architecture. In *Proc. Of Fusion 2003*, Cairns, Australia, July 2003.

[4] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Francisco, 1988.

[5] R. G. Cowell, A. P. Dawid, S. L. Lauritzen, and D. J. Spiegelhalter. *Probabilistic Networks and Expert Systems*. Springer-Verlag, New York, 1999.

[6] S-L. Normand and D Tritchler. Parameter updating in a bayes network. *Journal of the American Statistical Association*, Vol. 87, No. 420:1109–1115, December 1992.

[7] S. L. Lauritzen. Propagation of probabilities, means, and variances in mixed graphical association models. *Journal of the American Statistical Association*, Vol. 87, No. 420:1098–1108, December 1992.

[8] M. Paskin. Thin junction tree filters for simultaneous localization and mapping. In *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence (IJCAI-03)*, page pp. 11571164, San Francisco, CA, 2003.

[9] C. Berge. *Graphs and Hypergraphs*. American Elsevier Pub. Co., New York, 1973.

[10] B. D. O. Anderson and J. B. Moore. *Optimal Filtering*. Prentice-Hall, 1979.