OPTIMAL DATA INCEST REMOVAL IN BAYESIAN DECENTRALIZED ESTIMATION OVER A SENSOR NETWORK

T. Bréhard & V. Krishnamurthy

Dept. Electrical and Computer Engineering
University of British Columbia
2332 Main Mall,Vancouver, BC V6T 1Z4
e-mail: {brehard,vikramk}@ece.ubc.ca

ABSTRACT

A fundamental issue in Bayesian decentralized estimation over a sensor network is the inadvertent multiple re-use of data also known as *data incest*. We show the relationship between data incest and the network topology by using a graph theoretical formulation. A novel necessary and sufficient condition based on the topology of the network is derived so that data incest management can be optimally achieved. This approach requires large storage capabilities at the sensor level. In the case of an arbitrary network, if the necessary and sufficient condition for data incest does not hold then finding a sub-optimal strategy requires solving a 0-1 integer optimization problem where the dimension of the vector to optimize increases with time. Numerical results illustrate the effectiveness of our approach.

Index Terms— Sensor Network, Decentralized Estimation, Data Incest, Graph Theory

I. INTRODUCTION

Consider a network composed of multiple sensors used to collect noise corrupted observations of a target or situation. A centralized architecture which consists of each sensor transmitting measurements directly to a fusion center is theoretically optimal. However, this centralized approach has several disadvantages. First, a centralized architecture requires a high bandwidth to collect measurements and the fusion center requires high computation abilities. Second, survivability of the fusion center is critical. Finally, the user access to fused estimates can be slow. Thus, there is a great deal of interest in network-centric warfare applications [1] for decentralized architectures [2]. In a decentralized architecture, all information is processed locally at each sensor and no central processing is used. Due to bandwidth constraints, the sensors only communicate their estimates over the network and are not allowed to transmit raw measurements. One of the main issues in a decentralized architecture is to develop an algorithm so that each sensor is able to compute an estimate based on its own information and the estimates received from other sensors.

An estimate computed must have two properties. First, an estimate must be *optimal* in the sense that it is equal to the estimate obtained in an *ideal network* where measurements instead of estimates are broadcast through the network. Second, an estimate must be *free of data incest*. The recursive nature of the estimation process requires to pay attention to the possible abusive re-use of measurement information [3].

The key requirement is to fuse estimates sharing a common information set. A first approach consists of formulating this problem as fusion with unknown correlation. A consistent estimate can be derived using Covariance Intersection [4]. However, this algorithm based on the convex combination of information is by nature sub-optimal. The recursive nature of the estimation process implies the knowledge of the information already integrated. A second approach is to use consensus avering technique based on an asymptotical property [5]. A third approach consists of assuming a complete or partial knowledge of sensor network topology. Using this kind of assumption, optimal estimates free of data incest can be derived for some particular topologies. An optimal solution can be computed for fully interconnected networks using the Decentralized Kalman Filter [6] or the Decentralized Information Filter [7]. Grime et al [7] solved the case of connected tree networks by combining a Decentralized Information Filter and a Channel Filter. Dodin et al extended this principle in [8] to geodetic graphs. McLaughlin [3] investigated the case of Fully Interconnected Networks which Maximum communication delay using a Decentralized Information Filter combined with a Data Incest Management. He derived a sufficient condition on the topology of the network which guarantees that data incest problem can be solved. Chong et al introduced graph information which is a directed graph to identify common information in [9].

In this paper, we investigate the case of arbitrary network topologies using graph theory. This approach is based on the work of McLaughlin [3] which showed that information flow in the network can be represented by a Directed Acyclic Graph. First, we show in Section II that the *Adjacency Matrix* and the *Transitive Closure Matrix* which are classical tools of graph theory are the key tools to understand the topology of the network and consequently to derive optimal estimates free of data incest. Then, an original necessary and sufficient condition on the topology of the network which guarantees that optimal estimates free of data incest can be computed is derived. We show in Section III that if this condition does not hold, finding a sub-optimal strategy implies solving a 0-1 integer optimization problem. This approach is illustrated by simulation results in Section IV.

II. DECENTRALIZED ESTIMATION IN AN ARBITRARY NETWORK

Consider a network consisting of S sensors where each sensor observes an unknown d-dimensional vector x. Let $z_{\lfloor k,s \rfloor}$ be the measurement observed by sensor s at time k:

$$z_{\lfloor s,k \rfloor} = H_{\lfloor s,k \rfloor} x + v_{\lfloor s,k \rfloor} \ , s \in \{1, \dots, S\}, \ k \in \mathbb{N}^* \qquad (1)$$

where $v_{\lfloor s,k \rfloor}$ denotes a white Gaussian noise sequence with covariance matrix $R_{\lfloor s,k \rfloor}$. For the sake of simplicity, the following re-indexing scheme is used:

$$n \triangleq \lfloor s, k \rfloor \triangleq s + S(k-1) \; . \tag{2}$$

Consequently, (1) can be rewritten:

$$z_n = H_n x + v_n \; , \tag{3}$$

H_n is a known matrix that relates the unknown parameter to measurement n. At each time instant k, each sensor must update its estimate using the measurement observed and the information received from other sensors. Let us consider two types of networks.

Definition 1 (Ideal network): *In an ideal network, the sensor nodes broadcast all the measurements they have observed or received.*

Definition 2 (Constraint network): *In a constraint network, the sensor nodes broadcast their estimates.*

An ideal network is inconvenient in practice as the amount of information to transmit is too large. However, this network is interesting from a theoretical point of view. As shown in section II-B, estimates free of data incest can always be computed in an ideal network. Moreover, these estimates are directly related to some classical tools of graph theory discussed in section II-A. We derive in section II-C an original necessary and sufficient condition which guarantees that the estimates in constraint network and ideal network will be equal.

II-A. Decentralized estimation and graph theory

We review the definitions of the Adjacency Matrix and the Transitive Closure Matrix which are classical tools of graph theory to study the communication topology of graph. We will see in section II-B that the optimal estimates free of data incest in ideal networks are directly related to these two matrices.

Definition 3 (Graph): *A graph G_n is a pair (V_n, E_n) where $V_n \triangleq \{v_1, \ldots, v_n\}$ is the set of vertices and $E_n \subset V_n \times V_n$ the set of edges between the vertices.*

Definition 4 (Adjacency Matrix): *Let $G_n = (V_n, E_n)$ be a graph. The Adjacency Matrix A_n of G_n is a $n \times n-$matrix whose entries, $A_n(i,j)$ are given by:*

$$A_n(i,j) = \begin{cases} 1 & \text{if } (v_j, v_i) \in E_n \; , \\ 0 & \text{otherwise} \end{cases} \; . \tag{4}$$

Definition 5 (Transitive Closure Matrix): *Let $G_n = (V_n, E_n)$ be a graph. The Transitive Closure Matrix $\mathcal{T}_n$ of G_n is a $n \times n-$matrix whose entries, $\mathcal{T}_n(i,j)$ are given by:*

$$\mathcal{T}_n(i,j) = \begin{cases} 1 & \text{if there is a path beween } v_j \text{ and } v_i \; , \\ 0 & \text{otherwise} \end{cases} \; . \tag{5}$$

The information flow in a sensor network can be represented by a Directed Acyclic Graph (DAG). A node $v_j \in V_n$ represents a specific sensor s_j at a specific time k_j. This choice implies that the set of edges does not represent only the transmission of information between different sensors. There is an edge from vertex v_j to vertex v_i in three cases:

1) an information is broadcasted from sensor s_j at time k_j and received by sensor s_i at time k_i ,
2) $s_j = s_i$ and $k_j < k_i$,
3) $s_j \neq s_i$ and it exists $l \neq i$ such that $s_l = s_i$, $k_l \leq k_i$ and $(v_j, v_l) \in E_n$.

The second case arises when two vertices of the graph correspond to the same sensor at different time. If an information is available at time k_j, it is still available at time k_i. The third case corresponds to the case where a vertex s_i received an information from a different sensor s_j at a previous time k_l. Now, using the re-indexing scheme (2), we deduce that if $i < j$ then $(v_j, v_i) \notin E_n$. Consequently, the Adjacency Matrix is a strictly upper triangular matrix and G_n is a DAG. Then, the Transitive Closure Matrix $\mathcal{T}_n$ is directly computed from A_n using the following formula:

$$\mathcal{T}_n = g(\{\mathcal{I}_n - A_n\}^{-1}) \; , \tag{6}$$

where $\mathcal{I}_n$ is the $n \times n$ identity matrix. g is a function such that for a $n \times n$-matrix B, $C \triangleq g(B)$ is a $n \times n$-matrix where $C(i,j)$ is equal to zero if $B(i,j) = 0$ and one else. Eq.(6) is derived from the classical interpretation of matrix $\{\mathcal{I}_n - A_n\}^{-1}$. Entry in row i and column j of this matrix gives the number of paths from vertex i to vertex j.

II-B. Optimal estimation for ideal networks

We derive new formulas to compute optimal estimates free of data incest in an ideal network at node n in this section. The optimal estimate is presented in proposition 1 using the classical information filter notations which are reproduced below for convenience:

$$\begin{cases} i_n \triangleq H_n^t R_n^{-1} z_n \; , \\ I_n \triangleq H_n^t R_n^{-1} H_n \end{cases} \text{and} \begin{cases} \hat{y}_n \triangleq \hat{P}_n^{-1} \hat{x}_n \\ \hat{Y}_n \triangleq \hat{P}_n^{-1} \; , \end{cases} \tag{7}$$

where R_n is the measurement error covariance at node n and H_n is defined by eq.(3). We note $\hat{x}_n$ the optimal estimate observed by node n, and $\hat{P}_n$ the estimation error covariance matrix. We also introduce the following vector notation where ^{t} denotes the transpose:

$$i_{1:n} \triangleq [i_1^t, \ldots, i_n^t]^t \; , \text{ and } I_{1:n} \triangleq [I_1, \ldots, I_n]^t \; . \tag{8}$$

Proposition 1 (Optimal estimation in an ideal network): *Let G_n be a DAG representing an ideal sensor network. The optimal estimate free of data incest at node n is given by the following equations:*

$$\begin{cases} \hat{y}_n = (t_n \otimes \mathcal{I}_d) i_{1:n-1} + i_n \; , \\ \hat{Y}_n = (t_n \otimes \mathcal{I}_d) I_{1:n-1} + I_n \; , \end{cases} \tag{9}$$

where t_n is the $1 \times (n-1)$ left lower sub-matrix of T_n (the Transitive Closure Matrix). $\otimes$ is the direct matrix product and $\mathcal{I}_d$ is the $d \times d$ identity matrix where d is the dimension of the unknown state x.

Proof: Proof is presented in Appendix A. □

According to proposition 1, the optimal estimate free of data incest can be expressed as a linear function of the measurements via the transpose Transitive Closure Matrix T_n. This formula is quite intuitive, the optimal estimate at node n is the sum of the information collected by the nodes such that there are paths between all these nodes and n.

II-C. Optimal estimation for constraint networks

Based on the result on the optimal estimation in ideal networks, an original proposition for the optimal estimation in a constraint network is derived. We prove that estimates in the constraint network are equal to the estimates in ideal networks if and only if the topology of the graph as a topology verifying property 1 defined below. We summarize in Algorithm 1 the steps of the algorithm.

First, let us describe the class of estimates which can be built in the constraint network. At node n, the estimate is a weighted sum of estimates received from previous nodes $\tilde{y}_{1:n-1}$ and the current information i_n so that the estimate at node n has the following formulation:

$$\begin{cases} \tilde{y}_n = (w_n \otimes \mathcal{I}_d) \tilde{y}_{1:n-1} + i_n \; , \\ \tilde{Y}_n = (w_n \otimes \mathcal{I}_d) \tilde{Y}_{1:n-1} + I_n \; . \end{cases} \tag{10}$$

Moreover, we have a constraint w_n which indicates that some estimates are not available to node n due to the graph topology.

Constraint 1 (Topological constraint):

$$\text{if } a_n(j) = 0 \quad \text{then } w_n(j) = 0 \quad \forall\, j \in \{1, \ldots, n-1\} \tag{11}$$

where a_n is the $1\times(n-1)$ right upper matrix of adjacency matrix A_n associated to G_n. $\otimes$ is the matrix direct product. $\mathcal{I}_d$ is the $d \times d$ identity matrix and d the size of the unknown state x.

proposition 2 guarantees that there is a unique solution for w_n if the Adjacency Matrix A_n verifies property 1.

Property 1 (Optimality of graph): *Let G_n be a DAG. This graph has an optimal topology at node n if*

$$(a_n(j) - 1)w_n(j) = 0 \quad \forall\, j \in \{1, \ldots, n-1\} \tag{12}$$

where

- *a_n is the $1\times(n-1)$ right upper matrix of A_n the Adjacency Matrix ,*
- *w_n is the $1 \times (n-1)$ left upper matrix of T_n^{-1} (T_n is the Transitive Closure Matrix transpose).*

Proposition 2 (Optimal estimation in a constraint network): *Let G_n be a DAG representing a constraint sensor network. It is assume that the $n-1$ latest nodes have been able to compute optimal estimates. If and only if G_n satisfies property 1, the optimal estimate free of data incest at node n is given by the following formula:*

$$\begin{cases} \tilde{y}_n = (w_n^{opt} \otimes \mathcal{I}_d)\tilde{y}_{1:n-1} + i_n \ , \\ \tilde{Y}_n = (w_n^{opt} \otimes \mathcal{I}_d)\tilde{Y}_{1:n-1} + I_n \end{cases} \tag{13}$$

where vector w_n^{opt} is given by:

$$w_n^{opt} = t_n T_{n-1}^{-1} \tag{14}$$

and t_n is the $1\times(n-1)$ left lower sub-matrix of T_n (the Transitive Closure Matrix transpose). $\otimes$ is the matrix direct product. $\mathcal{I}_d$ is the $d \times d$ identity matrix and d the size of the unknown state x.

Proof: Proof is presented in Appendix A. □

By proposition 2, the existence of an optimal estimate for node n is directly related to the topology of the network. However, there are three major drawbacks to this approach:

1) We assume that a node knows exactly the communication topology of the network. This assumption may not be quite realistic, however, this is a first step toward a estimation process free of data incest based on a partial knowledge of the topology of the network.
2) If the network topology does not satisfy property 1, a sub-optimal strategy must be derived.
3) Equation (13) implies that node n has to store all the estimates computed through time and all the estimates received. This amount of information corresponds at most $n-1$ pieces of information. Consequently, the quantity of information to store increases with time.

III. SUB-OPTIMAL STRATEGY FOR ARBITRARY NETWORKS

According to proposition 2, if property 1 does not hold, an optimal estimate free of data incest can not be computed. We show that a sub-optimal strategy is the solution of a 0-1 integer optimization problem.

We assume that until node $n-1$, estimates free of data incest can be computed but that the assumption on network topology is not checked at node n. Then, optimal weight vector of proposition 2 can not be used because it involves unavailable information. A sub-optimal weight vector w_n^{sub} which verifies property 1 must be computed. The latter must be such that the sub-optimal estimate y_n^{sub} is free of data incest and closed to the optimal estimate $\hat{y}_n$. If we choose w_n^{sub} in the following set:

$$\begin{aligned} \mathcal{W}_n \triangleq \quad & \{w_n^* | w_n^* = t_n^* T_{n-1}^{-1}, t_n^* \in \{0,1\}^{\otimes n-1}\} \\ & \cap \{w_n^* | w_n^* \text{ verifies constraint 1 }\} \ , \end{aligned} \tag{15}$$

the estimate obtained is free of data incest and the weight vector verifies property 1. We now show why these weight vectors guarantees that the estimate compute is free of data incest. Let $\tilde{y}_n$ be the estimate computed using a weight vector from (15). Using the definition of w_n^* given by (15) and the formula of optimal estimate free of incest in ideal networks given by (9), eq.(10) becomes:

$$\tilde{y}_n = (t_n^* \otimes \mathcal{I}_d) i_{1:n-1} + i_n \ . \tag{16}$$

By definition t_n^* belongs to $\{0,1\}^{\otimes n-1}$. Consequently, each measurement information appears only once in the formula (16) and the estimate is free of data incest. Now, we would like to choose $w_k^* \in \mathcal{W}_n$ so that $\tilde{y}_n$ is closed to the optimal estimate free of data incest. This is a problem of optimization which can be written as follows:

$$w_n^{sub} = \underset{w_n^* \in \mathcal{W}_n}{\operatorname{argmin}} \|w_n^* - w_n^{opt}\| \ . \tag{17}$$

where $\mathcal{W}_n$ is defined by (15). Using the definition of w_n^* and w_n^{opt}, this problem can be rewritten as follows:

$$t_n^{sub} = \underset{t_n^* \in \mathcal{X}_n}{\operatorname{argmin}} \|t_n^* - t_n\| \ \text{ where} \tag{18}$$

$\mathcal{X}_n \triangleq \{t_n^* \in \{0,1\}^{\otimes n-1} | w_n^* = t_n^* T_{n-1}^{-1}$ verifies constraint 1 $\}$. This is a 0-1 integer optimization problem. Unfortunately, it seems that this problem does not have a closed-form solution in the general case. Moreover, as the dimension of the vector to optimize (i.e. $n-1$) is increasing with the number of nodes, this problem will quickly become intractable.

IV. SIMULATIONS RESULTS

To illustrate the performances of the optimal estimate free of data incest (algorithm 1) in the context of an arbitrary network, a network of 20 sensors having a ring communication topology with a random delay is considered. The communication delay is the same for each of the sensors but changes randomly through time. One can show using simulation that this topology verifies property 1. Then, according to proposition 1, the optimal estimate free of data incest (algorithm 1) can be computed. The latter is compared with the classical Covariance Intersection (CI) [4]. The comparison is based on three criteria:

- Mean Square Error (MSE),
- Error ellipse area,
- Storage requirements.

The duration of the scenario is 100 seconds. The measurement error covariance associated to measurement $z_{\lfloor s,k \rfloor}$ is

$$R_{\lfloor s,k \rfloor} = \begin{pmatrix} \sigma_x^2 \sin^2\beta_s + \sigma_y^2 \cos^2\beta_s & (\sigma_y^2 - \sigma_x^2)\sin\beta_s\cos\beta_s \\ (\sigma_y^2 - \sigma_x^2)\sin\beta_s\cos\beta_s & \sigma_x^2 \sin^2\beta_s + \sigma_y^2 \cos^2\beta_s \end{pmatrix} \tag{19}$$

where $\beta_s = \operatorname{atan}(r_y^s - r_y, r_x^s - r_x)$. Terms (r_x^s, r_y^s) and (r_x, r_y) are respectively the position of sensor s and the position of the target on x-y plan. Constants σ_x and σ_y are fixed to $8\ m$ and $40\ m$. The measurement covariance matrix (19) is such that each sensor gives a *"specific information on the target position"* which depends on relative position of the target to the sensor.

We can see in Fig.1 that the MSE associated to the optimal estimate free of data incest is lower than the MSE associated to the CI. Fig.2 shows the error ellipse area for the optimal estimate free of data incest versus CI. The error ellipse area for the optimal estimate free of data incest is lower than error ellipse area for CI. Moreover, if we compare the performance for the last node at the end of scenario, we observe that the optimal estimatation free of data incest is $[150, 150] \pm 0.75\ m$ when the CI gives $[150, 150] \pm 15\ m$. Consequently, the performances of the new algorithm in terms of accuracy and confidence are better. Fig.3 presents the variation of the quantity of information to store locally for each of two algorithms. The storage requirements for CI is constant through time when the storage requirement of the Optimal estimate free of data incest increases linearly with time. This fact has been explained in section II.

We have shown through simulation results that the optimal estimate free of data incest outperforms CI. These results were predictable because of the optimal nature of the new algorithm. However, the storage requirements of this new algorithm increase linearly with time. This drawback will be an important issue for future works.

V. CONCLUSION

We highlight in this paper the relationship between decentralized estimation networks and some classical tools of graph theory. The main result is an original necessary and sufficient condition based on the topology of the network which guarantees the optimality of the estimates. The algorithm outperforms the classical Covariance Intersection. This original framework based on graph theory offers a new way of regarding decentralized estimation networks problems. Future works will aim at enriching this framework to handle the estimation of Markovian processes and partially known communication topologies. Moreover, sub-optimal techniques will be developed to reduce the storage requirements.

APPENDIX A: PROOF OF PROPOSITION 1

First, left multiple eq.(3) by $(t_n(j) \otimes \mathcal{I}_d)H_j^t R_j^{-1}$ to obtain $\hat{x}_n$ mean square estimate and $\hat{P}_n$ the estimation covariance matrix:

$$\begin{cases} \hat{x}_n = \hat{P}_n \sum_{j=1}^{n} (t_n(j) \otimes \mathcal{I}_d) H_j^t R_j^{-1} z_j \ , \\ \hat{P}_n^{-1} = \sum_{j=1}^{n} (t_n(j) \otimes \mathcal{I}_d) H_j^t R_j^{-1} H_j \end{cases} . \quad (20)$$

Then, express eq.(20) using the information filter notations (see eq.(7)).

APPENDIX B: PROOF OF PROPOSITION 2

Using proposition 1 and remarking that $T_n(n, n)$ is always equal to one (see eq.(6)), eq.(10) gives the following equality:

$$(t_n \otimes \mathcal{I}_d) i_{1:n-1} = (w_n^{opt} \otimes Id_d)(T_{n-1}^t \otimes \mathcal{I}_d) i_{1:n-1} \ . \quad (21)$$

This relation must be true for any measurement information so that w_n^{opt} formula given by eq.(14) is proved. Now remind that T_n is a upper triangular matrix with ones on the diagonal, consequently w_n^{opt} is the $1 \times (n-1)$ right upper matrix of T_n^{-1}. Then, w_n^{opt} will verify constraint 1 if and only if property 1 is true.

VI. REFERENCES

[1] L.J. Sciacca and R.J. Evans. Cooperative Sensor Netwoks with Bandwith Constraints. In *Proc.* SPIE *AeroSense Conf., Orlando, FL*, 1997.

[2] M.E. Liggins, C.-Y. Chong, I. Kadar, M.G. Alford, V. Vannicola, and S. Thomopoulos. Distributed Fusion Architectures and Algorithms for Target Tracking. In *IEEE*, volume 85, 1997.

[3] S.P. McLaughlin, R.J. Evans, and V. Krishnamurthy. A Graph Theoretic Approach to Data Incest Management in Network Centric Warefare. In *Proc. 8th Int. Conf. on Information Fusion, Philadelphia, PA*, 2005.

[4] S.J. Julier and J.K. Uhlmann. A Non-divergent Estimation Algorithm in the Presence of Unknown correlations. In *Proc. Amer. Contr. Conf., Albuquerque, NM*, 1997.

[5] L. Xiao, S. Boyd, and S. Lall. A scheme for robust distributed sensor fusion based on average cosensus. In *In Proc. of the 4th Int. Symp. on Info. Proc. in Sensor Net.*, 2005.

[6] H.F. Durrant-Whyte, B.Y.S. Rao, and H. Hu. Toward a Fully Decentralized Architecture for Multi-Sensor Data Fusion. In IEEE *Int. Conf. Robotics and Automation*, 1990.

[7] S. Grime and H.F. Durrant-Whyte. Communication in Decentralized Systems. In IFAC *Control Engineering Practice*, 1994.

[8] P. Dodin and V. Nimier. Distributed Tracking Systems and Their Optimal Inference Topology. In *Proc. 5th Int. Conf. on Information Fusion, Annapolis, MD*, 2002.

[9] C.-Y. Chong, S.Mori, and K.-C. Chang. Distributed Multitarget Multisensor Tracking. *in "Multitarget Multisensor Tracking", Artech House*, pages 247–295, 1990.

Algorithm 1 Optimal estimation in constraint network

For $k = 1, \ldots$ and $s = 1, \ldots, S$

1) Compute the optimal weight vector w_n^{opt} using eq.(14)
2) Compute the estimate $\tilde{y}_n$ and $\tilde{Y}_n$ using eq.(13)
3) Compute the estimate $\tilde{x}_n = \tilde{Y}_n^{-1}\tilde{y}_n$ and $\tilde{P}_n = \tilde{Y}_n^{-1}$

End for

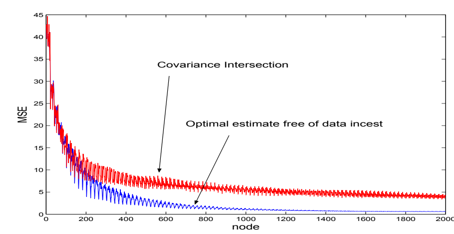

Fig. 1. Mean Square Error for Covariance Intersection versus optimal estimate free of data incest.

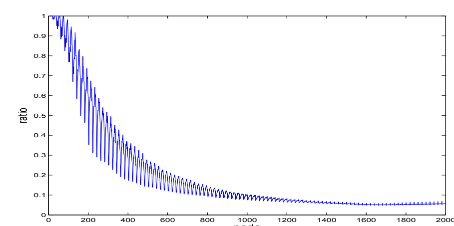

Fig. 2. Ratio of error ellipses area for optimal estimate free of data incest versus Covariance Intersection.

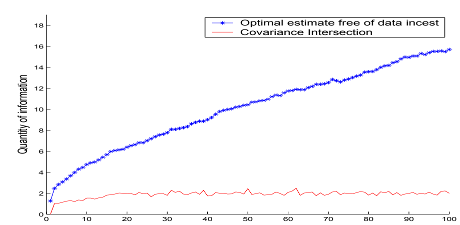

Fig. 3. Variation of the quantity of information stored at sensor level through time for Covariance Intersection estimate versus optimal estimate free of data incest.