Horizontal Integration based upon Decentralized Data Fusion (DDF), NetCentric Architecture (NCA), and Analysis Collaboration Tools (ACT)

C4ISR/C2 Architecture

Sheldon B. Gardner
Naval Research Laboratory
Washington, DC 20375
202-767-5652 (v), 202-404-7234 (fax)
gardner2@ccs.nrl.navy.mil

Abstract

The major focus in the joint-services area today is on Horizontal Integration (HI)– rapidly fusing and exploiting the data from different collection systems to speed the flow of correlated intelligence to war fighters, both for situational awareness and targeting. In this paper we discuss several technologies potentially useful in HI. They are Decentralized Data Fusion (DDF), NetCentric Architecture (NCA), and Analysis Collaboration Tools (ACT). DDF, NCA, and ACT could provide pathfinders and enablers for future implementation of HI across current 'stovepipe' collection and analysis systems. HI will ultimately require effective command and control of air, ground, naval, and space-based intelligence collection and dissemination systems. This control will be achieved through net-centric information management that includes dynamic links between a 'global' database and multiple locally maintained databases that contain data obtained from component stovepipe systems. The links between communication nodes will allow global information to be updated automatically with information from distributed assets. The motivation for providing connectivity and automated information sharing among distributed platforms/nodes is to increase the amount of information available at each node. The technology enablers for HI include, but are not limited to, the following:

- **Net-Centric Architecture (NCA)**- A network architecture that gives component platforms access to multi-level security information and communications over a two-way encrypted TCP/IP connection allowing net-centric control and utilization of ISR assets.
- **Decentralized Data Fusion (DDF)**-The DDF framework includes a proposed solution to the data fusion problem called Covariance Intersection (CI), as well as a solution to the information corruption problem called Covariance Union (CU).
- **Analysis Collaboration Tool (ACT**)- NRL has developed an analysis collaboration tool (ACT) to be used for virtual collaboration in intelligence support. ACT is a second-generation tool for implementation of application sharing and collaboration between analysts. Second generation meaning that ACT can be used to collaborate complex analysis applications ‘out of the box’ without source code changes or the need for configuration management.

1. Introduction

In the past decade changes in the threat environment require a new operating model which integrates data from different intelligence sources and service components. This model has been called Horizontal Integration (HI) since it represents a shift from vertical systems (e.g. stovepipes), structured to support specific needs, to a horizontally integrated enterprise structured across multiple disciplines. Current vertical systems call for actions to be performed by relatively autonomous functional nodes. In an HI environment these nodes are required to communicate via information networks. Each functional node incorporates the data it needs to perform its function and then processes and transmits its best available information for use by other nodes. In some cases the information produced by a component node is unique. However, in many cases, information is redundant and must be fused to produce an improved estimate. A current goal of HI is to increase the fidelity and availability of intelligence data. This will be achieved in part through a network-centric architecture (NCA) that will include dynamic links between a 'global' database and multiple locally maintained databases that could represent data obtained from component stovepipe systems. These links will allow global information to be updated automatically with information from distributed assets. For example, in carrier battle group operations, local track maintenance systems have moved in a similar direction with the evolution of multi-platform distributed sensing and tracking architectures that allow, multiple carrier group platforms to coordinate sensing and tracking assets in many-on-many engagement scenarios. The intuitive motivation for providing connectivity and automated information sharing among distributed platforms/nodes is to increase the amount of information available at each node. [1]

Effective HI will require an established framework for exploiting the benefits of decentralized data fusion (DDF) that ensures the evolution of consistent actionable intelligence among communicating nodes. Achieving this goal requires a solution to two major problems: (1) Invalid incorporation of redundant information and (2) Information corruption. A unified solution for DDF that ensures information consistency in the HI environment is required. At present, there is no established framework for exploiting DDF. In Section 2 we propose a unified solution for DDF that ensures information consistency in the HI environment. This DDF framework includes a solution to the data fusion problem called Covariance Intersection (CI), and a solution to the information corruption problem called Covariance Union (CU). Finally, in Section 3, we discuss the technology required to achieve the effective human-to-human collaboration needed for HI. We discuss the use of a software tool, developed by NRL, called Analysis Collaboration Tool (ACT). ACT can be used to share and collaborate highly complex analysis tools 'out of the box' without source code changes, added software, or configuration management. ACT is also cross-platform meaning that full collaboration can be achieved between Intel-based, Sun, and SGI workstations.

2.1 Decentralized Data Fusion with Redundant Information: Covariance Intersection (CI)

The most serious problem arising in DDF is the effect of redundant information. Specifically, pieces of information from multiple sources cannot be combined using traditional methods unless they are independent or have a known degree of correlation. It has been shown that maintaining correlation information in a fully general distributed processing network is computationally intractable. Specifically, currently proposed systems call for complex tasks to be performed by a large number of relatively autonomous functional nodes, which communicate via a complex information network. Each functional node retrieves (or is pushed) the information it needs to perform its function and then processes and transmits its best available information for use by other nodes. The key requirements for information to be effectively processed in a decentralized system are:

1. The information (reports) must have a well-defined measure of uncertainty and confidence.
2. The fusion process must ensure that the database updates maintained by all entities are consistent.
3. The data fusion method must be robust to failures in the network caused by, for example, communications disruptions.
4. The connectivity of the nodes can be dynamically changed.
5. The data fusion framework must be efficiently scalable, e.g., to networks having many nodes.

For simplicity, the examples given in this paper will relate to combat systems and battlespace entities. The context can be transferred directly to intelligence product related counterparts. Virtually all real-world systems for tracking targets, e.g., aircraft or missiles from sequences of sensor measurements, represent information in the form of a state (or mean) vector with an associated error covariance matrix. For example, the state of an aircraft can be represented as a vector, **a**, consisting of the aircraft's estimated mean position and velocity, e.g., $\mathbf{a} = [x, y, z, vx, vy, vz]^T$, and an error covariance matrix **A** that expresses the uncertainty associated with the estimated mean. If the expected error in the mean vector is **e**, then the error covariance matrix **A** is an estimate of the expected value/matrix of $\mathbf{ee}^T$, i.e., $E[\mathbf{ee}^T]$. The estimate is said to be consistent (or conservative) if and only if $\mathbf{A} \geq E[\mathbf{ee}^T]$, or equivalently, if $\mathbf{A} - E[\mathbf{ee}^T]$ is positive definite or semi definite. In words, an estimate (**a**, **A**) is said to be consistent if and only if its covariance **A** does not underestimate the actual expected squared error in its mean **a**.

In many modern defense systems, information about targets is maintained in a distributed, or ***network-centric***, collection of local tactical databases containing mean and covariance estimates of the states of relevant battlespace entities. So, continuing the example above, the estimate (**a**,**A**) relating to the state of an aircraft may be produced from a tracking system onboard a ship in a carrier group while a different estimate (**b**, **B**) of the same aircraft is produced by an airborne tracking system. If these sea and air assets establish a communication link, they can in principle exchange their respective estimates of the state of the aircraft and fuse them to produce a better (i.e., smaller error covariance) estimate (**c**, **C**) that can be targeted with greater precision.

If the error covariances **A** and **B** derive from statistically independent error processes, i.e., the estimates are uncorrelated, then a Kalman filter can be applied to fuse the two estimates to produce the improved fused estimate (**c**, **C**). However, in virtually any practical application the estimates will contain correlated error components resulting from the use of imperfect kinematic and/or measurement models. These correlations undermine the integrity of ***any*** Kalman filter. This correlation problem plagued industrial and defense applications of the Kalman filter for over forty-five years!

At present, most types of operational intelligence information is maintained in a variety of forms for a variety of different uses (e.g., precision target tracking, weapon-target allocation), and most of it consists of measured or estimated states of entities of interest that include associated measures of uncertainty, e.g., in terms of confidence or geographical location uncertainty. For example, the position of a target observed by an unmanned aerial vehicle (UAV) is computed from the UAV's estimate of its own position as determined from GPS and onboard inertial systems and from the relative observed position of the target based on the range and bearing measurements of its sensor. All of these components have known/modeled degrees of uncertainty that contribute to the overall uncertainty associated with the target's estimated position. This uncertainty is typically depicted graphically as an error ellipse. Pieces of information that include measures of uncertainty are commonly referred to as ***state estimates***.

For example, consider a weapon system where estimates of the state of a target are maintained at two nodes that have access to common information sources, then the respective state estimates will be correlated to some extent. Assume that the cross covariance information required for a Kalman filter fusion of the estimates is not available. If statistical independence is incorrectly assumed, then the fused estimate will be inconsistent, i.e., its associated error covariance matrix will underestimate the magnitude of the uncertainty associated with the mean state estimate. An underestimated error covariance can lead to incorrect likelihood estimates for its destination and also affect weapon allocation. For example, if a target's state has a spuriously smaller error covariance, then a high-precision weapon may be misapplied.

In order to avoid the potentially disastrous consequences of redundant data when using traditional methods, it is necessary to maintain cross covariance information. Unfortunately, it has been proven that in arbitrary decentralized networks it is not possible to maintain consistent cross covariance information. It is only possible in a few special cases, such as tree and fully connected networks, to avoid the proliferation of redundant information. These special topologies, however, fail to provide the reliability advantage because the failure of a single node or link results in either a disconnected network or one, which is no longer able to avoid the effects of redundant information. More intuitively, it is the redundancy of information in a network that provides reliability, so if the difficulties with redundant information are avoided by eliminating redundancy, then reliability also will be eliminated.

The problems associated with data fusion using correlated information have prevented the full power of decentralized architectures from being realized. What is needed is a framework that is able to fuse correlated/redundant information in a robust manner.

Toward this end, a mathematical framework has been developed by Uhlmann and Julier, called Covariance Intersection (CI) that permits the consistent fusion of arbitrary pieces of information in a totally decentralized network [2, 3]. Specifically, given mean and covariance estimates $\{\mathbf{a},\mathbf{A}\}$ and $\{\mathbf{b},\mathbf{B}\}$, a provably consistent fused estimate $\{\mathbf{c},\mathbf{C}\}$ can be generated from the following CI equations:

$$\begin{aligned} \mathbf{C} &= (w\mathbf{A}^{-1} + (1-w)\mathbf{B}^{-1})^{-1} \\ \mathbf{c} &= \mathbf{C}(w\mathbf{A}^{-1}\mathbf{a} + (1-w)\mathbf{B}^{-1}\mathbf{b}), \end{aligned}$$

where w in the interval [0,1] is a parameter that determines which measure of covariance size is minimized (e.g., determinant, trace, etc.). The fused estimate is guaranteed to be consistent regardless of the degree of correlation between the estimates to be fused. Any choice of w is guaranteed to produce a consistent estimate such that the covariance $\mathbf{C}$ is greater than or equal to the true squared error in the mean estimate $\mathbf{c}$. This result circumvents the assumptions made in proofs that it is impossible to ensure nondivergence in decentralized networks. Thus, for the first time, it is possible to place distributed/decentralized data fusion on a rigorous mathematical foundation.

2.2. The Information Corruption Problem: Covariance Union (CU)

In the intelligence context, the potential benefits HI are clear, but there are also clear operational obstacles that must be surmounted. Covariance Intersection completely solves the most general form of the data fusion problem. In practice, however, a different problem can arise before data fusion can even be applied. In particular, what happens if a spurious or corrupted estimate enters the network? The answer is that the spurious estimate will corrupt every estimate with which it is fused and every estimate that is subsequently fused with any of the newly corrupted estimates. In other words, information in the network becomes corrupted at a geometric rate.

Specifically, what can be done if two estimates ($\mathbf{a}$, $\mathbf{A}$) and ($\mathbf{b}$, $\mathbf{B}$), purportedly relating to the state of the same real-world object, are statistically inconsistent with each other? For example, if two mean position estimates differ by more than a kilometer, but their respective covariances suggest that each mean is accurate to within a meter, then clearly something is wrong. Resolving such inconsistencies among estimates is sometimes referred to as ***database deconfliction***.

A statistical method, Covariance Union (CU), has been developed that addresses the database deconfliction problem for mean and covariance estimates [3]. Specifically, a set of CU operations can be applied in a fully automatic fashion at each node in a network to ensure that estimates that are in conflict with the consensus of information in the network will be completely filtered out or strongly limited in their ability to propagate. What is remarkable is that this capacity to eliminate spurious information is an emergent property of the locally applied operations - there is no centralized mechanism that tries to identify which pieces of information are spurious and which are not. The CU mechanism provides two absolutely critical features:

1. It provides a mathematically rigorous, yet computationally efficient, method for identifying pairs of estimates that are statistically inconsistent with each other according to a specified level of tolerance. The method alone cannot possibly identify which of the estimates is spurious/corrupted, but it provides a trigger for the application of the CU fusion operation.

2. The CU fusion operation replaces two inconsistent estimates with a single estimate that is statistically consistent with ***both*** of the given estimates. Such an estimate can be consistent with the two estimates only by having a sufficiently large associated

degree of uncertainty (i.e., large error covariance). The CU mechanism is optimal in the sense that it constructs the estimate with the smallest possible sufficiently large covariance. This estimate can be safely propagated and assimilated throughout the network. It has been proven that the local application of the CU mechanism at each node can eliminate spurious/corrupted estimates in almost all practical situations without in any way adversely affecting the converged operational picture information.

The Covariance Intersection method guarantees consistency as long as the system and measurement estimates are each consistent. In the deconfliction problem it is only known that one of the estimates, either (**a**, **A**) or (**b**, **B**), is a consistent estimate of the state of the object of interest. Because it is not generally possible to know which estimate is spurious, the only way to rigorously combine the estimates is to form a unioned estimate, (**u**, **U**), that is guaranteed to be consistent with respect to ***both*** of the two estimates. Such a unioned estimate can be constructed by applying convex or semidefinite optimization methods to find a mean vector **u** and covariance matrix **U** such that:

$$\mathbf{U} \geq \mathbf{A} + (\mathbf{u}-\mathbf{a})(\mathbf{u}-\mathbf{a})^T$$
$$\mathbf{U} \geq \mathbf{B} + (\mathbf{u}-\mathbf{b})(\mathbf{u}-\mathbf{b})^T$$

where some measure of the size of **U** (e.g., determinant) is minimized. This Covariance Union (CU) of the two estimates can be subsequently fused with other consistent estimates using CI.

Intuitively, the above equations say that if the estimate (**a**, **A**) is consistent, then the translation of the vector **a** to **u** will require its covariance to be enlarged by the addition of a matrix at least as large as the outer product of (**u**-**a**) in order to be consistent. The same reasoning applies if the estimate (**b**, **B**) is consistent. Covariance Union therefore determines the smallest covariance **U** that is large enough to guarantee consistency regardless of which of the two given estimates is consistent. As a simple 2D example, suppose **a**=[0,0], **b**=[4,4], and each estimate has an error covariance equal to the identity matrix **I**. If the two estimates are determined to be statistically inconsistent with each other, the optimal CU deconflicted estimate can be determined to be

$$\mathbf{u}=[2, 2], \quad \mathbf{U}= \begin{vmatrix} 5 & 4 \\ 4 & 5 \end{vmatrix}.$$

It is straightforward to verify that this estimate is in fact consistent with respect to both of the estimates: If (**a**, **A**) is the valid state estimate, then the covariance **U** for mean **u** must be greater than or equal to $\mathbf{A} + (\mathbf{u}-\mathbf{a})(\mathbf{u}-\mathbf{a})^T$, which it is. It can be verified that the estimate is also consistent with respect to (**b**, **B**). Therefore, if either of the two estimates is a consistent estimate of the state of the object of interest, then the CU estimate is also consistent.

More generally, for any set of *n* measurements in the same coordinate frame, e.g., ($\mathbf{a}_1$, $\mathbf{A}_1$), ($\mathbf{a}_2$,$\mathbf{A}_2$), ..., ($\mathbf{a}_n$, $\mathbf{A}_n$), in which one or more elements of the set is a measurement of a system of interest, it is possible to construct a unioned measurement (**u**, **U**) that is consistent with respect to each element of the set of measurements. In particular, (**u**, **U**) is defined by the following constraints:

$$\mathbf{U} \geq \mathbf{A_1} + (\mathbf{u}-\mathbf{a_1})(\mathbf{u}-\mathbf{a_1})^T$$
$$\mathbf{U} \geq \mathbf{A_2} + (\mathbf{u}-\mathbf{a_2})(\mathbf{u}-\mathbf{a_2})^T$$
$$\vdots$$
$$\mathbf{U} \geq \mathbf{A_n} + (\mathbf{u}-\mathbf{a_n})(\mathbf{u}-\mathbf{a_n})^T$$

This unioned estimate (**u**, **U**) is guaranteed to be consistent as long as at least one element of the set of measurements is consistent with respect to the state of the object of interest. Thus, CU effectively solves the database deconfliction for all cases except that in which every one of the measurements/estimates is spurious, and it can be argued that no solution can exist to cover that possibility because no information is actually available.

3. NRL Analysis Collaboration Tool (ACT)

DoD has established a standard set of collaboration tools designated as the Defense Collaboration Tool Suite (DCTS). The tools permitted for use in a collaborative (human-to-human interaction) environment are those that include Microsoft NetMeeting or Sun SunForum as the basic building blocks. DCTS also includes CUSEEME Networks, Meeting Point Servers, and Digital Dashboard. However, when intelligence analysts collaborate, it is usual that participants desire to use their favorite analysis tools. In most instances, these tools will not work within the framework of the current DCTS. This places a serious limitation on the value of real-time collaboration.

In recognition of this limitation, the Naval Research Laboratory (NRL) has developed an analysis collaboration tool called (ACT) to be used for virtual collaboration in HI support [5]. ACT is a second generation tool for implementation of collaboration between analysts. First generation collaboration tools, exemplified by applications such as InfoWorkSpace (IWS), NetMeeting, and SunForum, are limited by the complexity of the analysis tools which can be shared and collaborated. In contrast, ACT can be used to share and collaborate highly complex analysis tools ‘out of the box’ without source code changes or added software. ACT technology is also cross-platform meaning that collaboration can be between Intel, Sun, and SGI platforms.

ACT is designed to work with existing DCTS applications, such as NetMeeting, SunForum, or SGI Meeting. The collaborative session works as follows: ACT is installed and run on a Host prior to the session. The collaboration session is started on the Host using NetMeeting, SunForum, or SGI Meeting. The collaboration host runs and shares the target application to all participants. Any participant can then take control of the target application using a ‘pass-the-chalk’ paradigm. During a session, all data remains on the host. Only analysis results (graphs, spreadsheets, etc.) are sent over the network. Consequently, data transfer requirements are minimized enabling operations over limited bandwidth networks.

Summary

Achieving HI as an operational military capability is a long process. Critical to this process is the conduct of field experiments and demonstrations incorporating new technologies that enable ISR systems to fuse and exploit intelligence data. The bottom line will ultimately be significant improvements in the flow of intelligence information to the warfighter. As a small step towards this goal, in this paper we have discussed DDF, NCA, and ACT as technology enablers. We are hopeful that other organizations, with responsibilities for planning and conducting experiments/demonstrations, can use these technologies to help achieve a networked, information sharing military force across the battlespace.

References

[1] Network Centric Warfare, D. Alberts, J. Garstka, F. Stein, CCRP Publication Series, 1999.

[2] Handbook of Multisensor Data Fusion, 2001, edited by David Hall and James Llinas. Chapter 12, ***General Decentralized Data Fusion with Covariance Intersection (CI)***, Simon Julier and Jeffrey Uhlmann.

[3] "Covariance Consistency Methods for Fault-Tolerant Distributed Data Fusion", J. Uhlmann, Information Fusion 4 p. 201-215, 2003.

[4] "Scalable Data Fusion," D. Nicholson, C.M. Lloyd, S.J. Julier, and J.K. Uhlmann, ***ISIF/IEEE Proceedings of the 5th International Conference on Information Fusion***, Annapolis, MD, 2002.

[5] "The Role of Analysis Collaboration Technology in Intelligence Support" , S. Gardner, ***Proceedings, Phoenix Challenge 2004 Conference***, Monterey, CA, 2004.