Track Purity and Current Assignment Ratio for Target Tracking Evaluation

Erik Blasch
Defence R&D Canada-Valcartier
2459 Pie-XI Blvd. North
Québec City, QC G3J 1X5
erik.blasch@drdc-rddc.gc.ca

Pierre Valin
Defence R&D Canada-Valcartier
2459 Pie-XI Blvd. North
Québec City, QC G3J 1X5
Pierre.valin@drdc-rddc.gc.ca

***Abstract** – Performance evaluation of simultaneous tracking and identification (STID) systems consists of measures of performance, measures of effectiveness, and measures of force effectiveness. To investigate the capability of STID, we extend track purity to a current assignment ratio. Track purity determines the correctly associated measurements for a given scan where as the Current Assignment Ratio (CAR) determines the correct measurements for a given track. Using the CAR aids sensor management algorithms in determining which targets have robust features for target identification such that in clutter, the target can be tracked. The CAR enables effectiveness evaluation of STID systems for mission success. In the paper, we review the fusion performance evaluation literature, outline STID metrics, and demonstrate the use of the CAR in a scenario.*

***Keywords:** Performance Evaluation, JBPDAF, Current Assignment Ration, Track Purity*

1 Introduction

Seminal texts in data fusion [1], target tracking [2, 3], and information fusion [4, 5] have addressed performance evaluation; yet the community as a whole has yet to adopt standards from which systems are uniformly evaluated. For example, the performance evaluation of multitarget simultaneous tracking and identification (STID) poses challenges for situational awareness.[6] STID is a subset of information fusion, which includes filtering, estimation, and prediction of data as well as a derivative evaluation over the context of the operating conditions of the sensor, targets, and environment.

Initial work on tracking performance evaluation was focused on optimal methods [7] which do not necessarily hold in a dynamic environment. Key developments and methods have been postulated and evaluated by X. Li and Y. Bar-Shalom [8], K. C. Chang [9], and C-Y. Chong and S. Mori [10]. Each of these approaches offers insight into the problem by clarifying useful measures of performance [MOPs] over tracking methods. We revisit the work of the above authors by looking at *track purity* [11, 12] and extending the method for STID analysis by using the *current assignment ratio* (CAR) [13]. Even in the last year, prominent researchers are looking at track purity as a metric of interest [14], as opposed to *track lifetime* [15] to understand the capability of forming tracks from measurements in clutter.

Use of features, attributes, and categorical representations of targets has become more prominent as users (or analysts) desire to not only know *where* the target is, but *who* it is, and even more practical, *what* is the target's behavioral intention. Initial use of track and ID methods sought to recognize landmarks for navigation [16, 17] and distinguish targets from clutter [18, 19]. By simultaneously processing target identification (who) and target tracking (where) can have mutual benefits to both reasoning systems [20,21]. To illustrate how the ID information may help in data association, **Figure 1** illustrates the process of how a target-ID can refine the positional measurement to select the validated measurement from the cluttered measurements. Numerous approaches in joint tracking and recognition (category), classification (type), and identification (fingerprint) have been applied using emerging techniques such as *Joint-Belief Probabilistic Data Association* (JBPDA) algorithm [22], pose-aiding radar [23], DSmT [24], and particle filters [25]. Features from radar [26, 27], infrared [28], and hyperspectral [29], data have been assessed. Recently, J. Dezert combined the PCR method with an IMM [30]. A combined track and ID algorithm can improve track quality, mitigate clutter confusion, and enhance target ID. Inherent in the research is the capability to discern target type location from different vantage points as fused from distributed platforms which requires integrated methods with sensor management. [31]

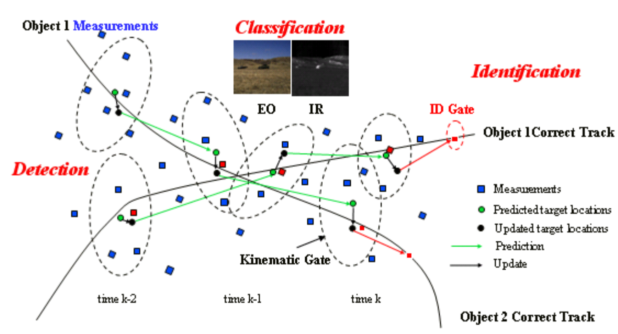

Figure 1. Data Association using ID and Position Measurements.

Sensor management includes the user placement and control of sensors, the automated processing, and the visualization of performance. [32] Information fusion systems must undergo rigorous tests before being operationally ready. Thus, routinely there are efforts to describe the objective and threshold metrics to guide testing of real-world systems. [33, 34, 35] To aid operational testing, numerous papers have tried to categorize the metrics of interest for not only the information fusion processing, but the system as a whole [36, 37, 38, 39, 40]. Various texts have provided measures of effectiveness in addition to the measures of performance. [1, 4]

The tracking community has supplied numerous papers and methods in *tracking* metrics and performance evaluation (PE). Initial work includes performance analysis of trackers in clutter [41], dense targets [42], and the algorithms themselves [43, 44]. As the need for track evaluation increased there were papers that summarized metrics [45] and new instantiations of some of the metrics [46]. K. C. Chang, S. Mori, and C-Y. Chong [47, 48, 49], continued to mature the techniques as well as X. R. Li and Z.-L. Zhao [50, 51, 52]. Numerous other examples exist of reporting results of research [53], tracking toolboxes [54], and most recently issues of computational costs and scalability [55].

Performance evaluation of *classification methods* is quite mature in the literature due to the elements of pattern recognition, image processing, and the security surveillance industry. Performance evaluation of classifiers typically includes receiver operator curves (ROCs) [56] that plot probability of detection versus probability of false alarms. Advances from the information fusion community include: confusion matrix analysis [57, 58], applications of Bayesian [59] and Dempster-Shafer methods [60], and on-line tools [61].

Similar to results from tracking and identification, there is a need for MOPs for situational awareness (SA). SA includes situational assessment and threat analysis with cues and inputs from users. Metrics and evaluation is not yet a well-established area of research, but can build from the developments of the tracking and classification communities. Examples include user involvement [62], situational awareness tools [63], and high-level MOEs [64]. Next we describe the measures of merit.

1.1 Measures of Merit

For complex Command and Control (C2) systems, the merit of the system can be established at various levels of observations. The Military Operations Research Society (MORS) developed a hierarchy of MOMs [Burton] for Command, Control and Information Systems (C2IS) that can be summarized as follows:

- Measures of Force Effectiveness (MOFEs): focus on how a force performs its mission or the degree to which it meets its objectives.
- MOEs: focus on the impact of C2 systems within the operational context.
- MOPs: focus on the internal system structure, characteristics and behaviour. MOPs of a system may be reduced to measures based on time, accuracy, capacity or a combination that may be interdependent.
- Dimensional Parameters: are the properties or characteristics inherent in the physical C2 systems.

Since the boundaries between the different levels can be quite fuzzy, this hierarchy provides rough divisions of a continuum of scales of observations, and serves as a guideline for the evaluation process. Some authors [Alberts] also suggest two other levels:

- The measure of *military utility*, tries to remove some of the scenario dependency of the measures.
- A measure of *policy effectiveness*, measures the worth of operations. Sometimes a successful mission is not a guarantee of overall success, i.e., winning a battle does not necessarily mean that the war will be won.
- A measure of *Command and Control effectiveness*, which measures the decision support capabilities.

Figure 2(a) shows the encircling relationship of the metrics and Figure 2(b) shows a balanced approach for metric analysis important to the user.

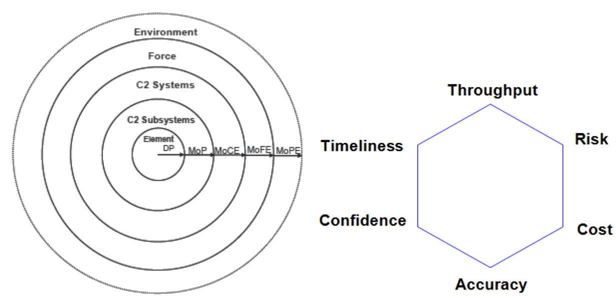

Figure 2. MOPs and MOEs relations.

Figure 3 highlights the summary from the *NATO Code of Best Practices for C2 Assessment* which highlights the importance of the metrics as well as the tradeoffs.

MoM	Focus	Scenario	Effort Required	Number	Impact	Compre-hension	Generaliz-ability
MoPE	Outcome	Dependent	High	Few	High	Policy	Low
MoFE	Mission						
MoCE	C3I	↕	↕	↕	↕	↕	↕
MoP	Systems						
DP	Process	Independent	Low	Many	Limited	Technical	High

Figure 3. MOM tradeoffs (from slides of Wallshein).

Figure 4 develops the MOMs in relation to the C3I system as a whole. Effectiveness is based on the function, structure and capability. The *NATO Code of Best Practices for C2 Assessment* also describes methods of evaluation through tests and scenarios of interest. The MOEs afford *speed* of data analysis, *efficiency* in communication, and *risk* reduction (or safety) from threats.

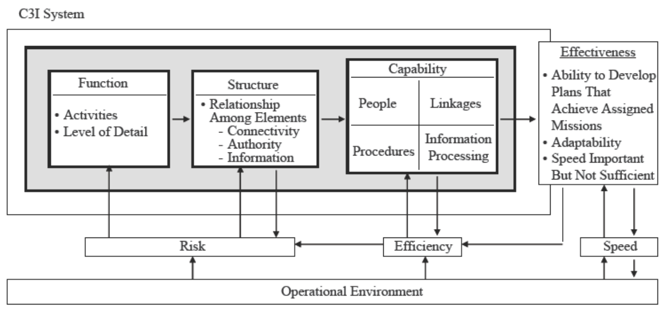

Figure 4. MOEs for C3I testing (from slides of Starr).

We have summarized key developments in fusion performance evaluation over tracking, classification, and system level analysis. Research over the entire spectrum would be longer than a paper; so we focus on one aspect in the scope of larger efforts. Here, we focus on extending the track purity MOP to afford MOE analysis for track and ID systems. Section 2 is a summary of performance evaluation with developments of track purity and the Current Assignment Ratio (CAR). Section 3 briefly describes the belief track and ID method. Section 4 presents a scenario to demonstrate the user of the CAR. Section 5 provides a conclusion and discussion.

2 Performance Evaluation

The recommended Measures of Performance (MOPs) quantify the following:

- Information accuracy: evaluates the quality of the positional tracking of the ground truth platforms in terms of the positional accuracy, the track purity and the correct assignment ratio.
- Information consistency: looks at the coherence in the information between a ship's database and the task coordinator database and the coherence between the organic and the non-organic data in the system. When a bad ID is associated to a track the inconsistency in the information can manifest itself as a track jumping. This type of inconsistency can be identified with track purity, correct assignment ratio and track continuity presented in other sections.
- Picture clarity: addresses the expected enhancement in terms of classification and identification of contacts as well as the system robustness to problematic sensor information generating false, redundant or spurious tracks.
- Picture completeness: evaluates how much of the real world the ship knows about. For the purpose of this analysis the real world is reduced to a specified region of space (the volume of interest, VOI) during a given time interval (the time interval of interest, TIOI).
- Track management statistics: permit an evaluation of how well the system behaves in real time. The load on the computer is measured in terms of the number of tracks and contacts is has to process and the time it takes to execute the different operations. Another question addressed by track management statistics is: how well handover of a track from a sensor to the other is performed in both systems. This is done by comparing the time and modality (manual, automatic) of track deletion with track continuity.

Some of these measures are performed on a single track (or ground truth platform) for an analysis of the stability of the information. Others are statistical measures based on many tracks/ground truth platforms and are used to establish an average behaviour of the system. A Measure of Force Effectiveness (MOFE), the model-based measure, was also proposed as an overall estimator of the system value.

A better understanding of the system performance will be gained if an effort is made to partition the different measures according to the type of tracks and region of interest (air, surface, underwater).

2.1 Track Purity and CAR

Track purity, a concept coined by Mori et al. [26], assesses the percent of correctly associated measurements in a given track, and so evaluates the association/tracking performance. This MOP is not explicitly dependent on detection performance, but it is dependent on the setting of association gates (which depends on the probability of detection P_d). It is also dependent on the ground truth platform density. Track purity measures the consistency with which a track is updated with measurements from a single ground truth platform or a set of ground truth platforms.

Correctional local MOPs, such as track purity, measure how well the tracks in an MSDF system are being associated with measurements of ground truth platforms. The track purity MOP is based on the calculation of a confusion matrix C for which the elements C_{ji} are constructed by counting reports. Given the tracks $t_1, \ldots, t_b$ and a set of ground truth platforms $g_1, \ldots, g_a$, C is:

Targets

$$
\begin{array}{c|cccc}
 & g_1 & g_2 & \cdots & g_a \\
\hline
 & C_{01} & C_{02} & \cdots & C_{0a} \\
t_1 & C_{11} & C_{12} & \cdots & C_{1a} \\
t_2 & C_{21} & C_{22} & \cdots & C_{2a} \\
\textbf{Tracks} & \vdots & \vdots & & \vdots \\
. & \vdots & \vdots & & \vdots \\
. & \vdots & \vdots & & \vdots \\
t_b & C_{b1} & C_{b2} & \cdots & C_{ba}
\end{array}
$$

Here, C_{ji} is the number of reports originating from ground truth platforms g_i which were assigned to track t_j *(i*

= *1, ..., a; j = 1, ..., b*) by the MSDF algorithm. Also, C_{0i} (the "ambiguity vector") consists of the number of reports that could not be assigned to any ground truth platform *(i = 1, ..., a)*. When C_{ji} is large, a strong association between t_j and g_i is implied.

Track purity is defined as the percentage of correctly associated measurements contained in a given track. The purity of the track t_j is defined as the normalised value of the largest element in the row defined by t_j:

$$TP\left[t_j\right] = \frac{\max\limits_{1\le i\le a} C_{ji}}{\sum\limits_{i=1}^{a} C_{ji}} \tag{1}$$

This measure can be estimated for each single track, but is more meaningful when statistics of this quantity are calculated. A recommended statistic is the Weighted Average of Track Purity (WATP) taken over all tracks and ground truth platforms. This statistic should be calculated separately for each type of track for air, surface and underwater platforms. It has a particularly convenient form if the weight given to each track is the number of measurements for that track, and if the weight given to each ground truth platform is the number of measurements originating from that ground truth platform. The resulting definition of the WATP is as follows:

$$WATP = \frac{\sum\limits_{j=1}^{b} \max_i C_{ji}}{\sum\limits_{j=1}^{b}\sum\limits_{i=1}^{a} C_{ji}} \tag{2}$$

The following elements are needed to compute Track Purity (TP) or WATP:

a. The list of correct (CO) track numbers for which TP will be computed (provided by the operator)
b. For each CO track pertaining to the selected CO track number, one needs the CO track number, the valid time and the ground truth platform number to which the CO track is attached
c. For each ground truth element corresponding to any ground truth platform number present in the selected CO tracks, one needs the time stamp and the ground truth platform number.

The confusion matrix is the starting point of many MOPs and its construction requires a lot a computation. Basically, we have to associate each CO Track report to a target in the ground truth. The choice of association can be made by a function of association that we will name *Associate*. This function will take as argument a track *T* at the time *t*, and the complete lists of tracks and ground truth's targets. *Associate* can be driven by positional and/or ID data. The resulting confusion matrix depends on the function *Associate* and it can be useful to test the related MOPs with some variations of *Associate*.

Here is a procedure to construct the confusion matrix that uses *Associate*:

a. Collect data to have all CO track reports for each track and each history point of all targets in the ground truth.
b. Initialize the confusion matrix by filling each entry with zeros.
c. For each track, process all CO track reports by:
   1) Using the association function *Associate*, to find the corresponding target in the ground truth
   2) Adding 1 to the related entry of the confusion matrix.

So the given algorithm can be automated if the association function *Associate* can be. The following paragraph discusses its feasibility.

First, the function *Associate* needs all the data aligned in time with the given CO track report. This can take a lot of computing time since it is proportional to the number of tracks and targets. This first step can be computed automatically without difficulty. The second step is to determine and use an association criterion that will select a target from the list of all targets in the ground truth. This step can also be automated since we can always select a target and, by hypothesis, the target list is not empty. The criterion can be based on position, ID, or both. Investigation has to be made to find the best criterion. Here, we will give some examples of criteria.

a. A rapid and easily implementable criterion is to choose the target that is the closest to this CO track. It is fast since it only proportional to the number of targets. However it can lead to erroneous results like associating all tracks with the same ground truth.
b. A better criterion is to use a Nearest Neighbour or a JVC association algorithm which requires more computation since it is proportional to the number of tracks and the number of targets. Based on position and/or ID, an intermediate association matrix has to be created to find the right association. Since each track belongs to a target in the ground truth, there is no problem when the number of tracks is lower or equal to the number of targets. The problem occurs when the number of tracks is greater than the number of targets (this may happen when a lot of spurious tracks are present). It would cause the algorithm to be unable to associate the CO track with a target.

The **Current Assignment Ratio** (CAR) [Gokberk] measures the performance for a ground truth platform instead of measuring the performance for a track. The CAR MOP for ground truth platform g_i is defined as the normalized value of the largest element in the column defined by g_i (i.e., by an analogous equation to TP, but maximising and summing over columns rather than rows). It assesses the percentage of contacts from a ground truth platform associated with the correct track.

$$CAR(g_i) = \frac{\max_{1 \le j \le b} C_{ji}}{\sum_{j=1}^{b} C_{ji}} \quad (3)$$

The higher the value of TP and CAR, the better the association performance is. Both measures are important and both should be measured, since the matrix C_{ji} has no special symmetry. Since the requirements to obtain this measure are the same as for TP, the difficulty and relevance to the present study are the same.

The CAR could be measured for a single ground truth platform, but a statistical quantity for many ground truth platforms is more significant. A useful statistic is the weighted average of CAR taken over all ground truth platforms (WACAR). It has a particularly convenient form if the weight given to each track is the number of measurements for that track, and if the weight given to each ground truth platform is the number of measurements originating from that ground truth platform. The resulting definition of the WACAR is as follows:

$$WACAR = \frac{\sum_{i=1}^{a} \max_j C_{ji}}{\sum_{i=1}^{a} \sum_{j=1}^{b} C_{ji}} \quad (4)$$

An example is presented in **Figure 4**:

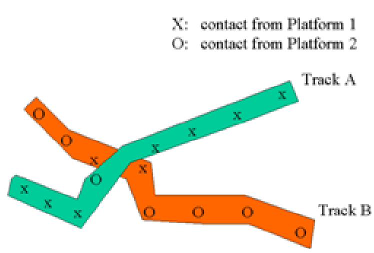

Figure 5. Example of contact to track Association

In this case the confusion matrix is:

	Platform 1	Platform 2
Track A	7	1
Track B	2	6

From this matrix the track purity is calculated as TP(A) = 7/8 = 87.5% and TP(B) = 6/8 = 75%and WATP = (7+6)/(7+1+2+6) = 13/16=81.3%. Similarly the correct assignment ratio is CAR(1)=7/9 = 77.8% and CAR(2) = 6/7 = 85.7% WACAR = (7+6)/(7+2+1+6)=13/16=81.3%.

2.2 Summary of Measures

The problem of track level and ID-level fusion has characteristic *tradeoffs* about which the sensor management system must arbitrate over a given scenario. Here we list some of the salient metrics.

Measures of Performance

1. *Information Accuracy*
   - Positional **Accuracy**
   - Track Purity
   - Correct Assignment Ratio
   - Accuracy of the filter covariance
2. *Information Consistency*
   - Proportion of target groups recognized
   - Proportion of recognized SA ROI in organic SA
3. *Information Currency*
   - Time in VOI Prior to Detection
   - Time from Detection to Confirmation
   - Data **throughput**
4. *Situational Clarity*
   - **Time** of Positive Classification / Positive Identification
   - Probability of detection and False Alarm Rate (FAR)
   - Target **Confidence**
   - Accuracy of Bayesian Percent Attribute Miss (BPAM)
   - Spurious Track Mean Ratio
   - Target Track Exchange Rate
5. *Situational Completeness*
   - Completeness History
   - **Value** – area coverage
6. *Track Management Statistics*
   - Time of track initiation / deletion
   - Track Continuity / Track Lifetime
   - Track swaps, broken tracks
   - Real-Time system Parameters

Measures of Effectiveness

1. Scenario Measures
   - **Timeliness** of information
   - **Survivability** as a function of detected targets
2. Threat evaluation
   - Percentage of Targets Correctly Assessed
   - Target Nomination Rate
   - Degree of Exactness in the Threat List Ranking
   - Information Gain
   - **Protection** (Protection = 1 – Threat Level)
2. Decision Support
   - Usability
   - Surveillance picture
   - **Safety** over all target position/IDs/intents (Safety = 1 – risk)

Measures of Force Effectiveness

1. Resource Management
   - Response Time over network communications
   - Time between target confirmation and weapon release
   - **Cost** of Battle over resources and people
2. Command-Level Support
   - Mission Analysis
   - **Interoperability** to send and receive contextual data

(see also the summary from Llinas, [4])

3 Track and ID Scenario

3.1 Problem Formulation

Consider an environment in which a tracker is monitoring multiple moving targets with stationary clutter. By assumption, the tracking sensor is able to detect target signatures. Assume that the 2-D region is composed of *T* targets with *f* features. Dynamic target measurements *z* are taken at time steps *k*, which include target kinematic and identification features $\mathbf{z}(k) = [x_t(k), f_1, \ldots f_n]$. A final decision from the STID algorithm is rendered as to which $[x, y]$ measurement is associated with the target-type.

The *multisensor-multitarget tracking and identification problem* is to determine which measured kinematic features should be associated with which ID features in order to optimize the probability that targets are tracked and identified correctly after *z* measurements. The multilevel feature fusion problem is formulated and solved by using concepts developed using the belief filter [18]. In the belief filter, the "association rule" uses the measurement with the highest target probability in a joint-belief probability data association (JBPDAF). The ID information is a result of the fused classification results where the aggregated classification is over 20 degree window pose measurements for the various targets.

3.2 Scenario and Track and ID Results

As detailed in the **Figure 6** below, by the true trajectory; the targets 1) start with position and velocity, 2) pass by each other at a close distance, and 3) finish with a specified direction. There was added noise to the true target position and clutter comprised of 5 spurious measurements around a target.

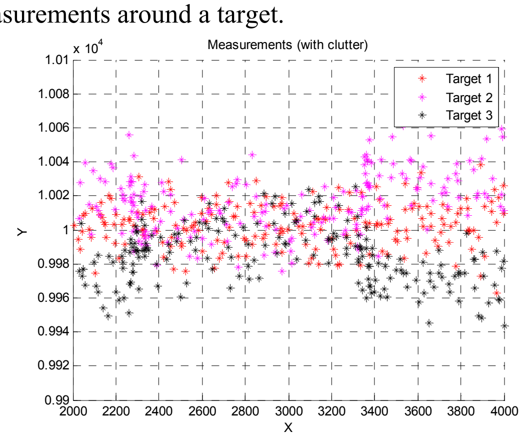

Figure 6. Track Scenario with Clutter Measurements.

Figure 7 shows the track only effectiveness when targets do not have ID information. The separation allows for the determination of a validation gate size that associates the correct measurements to tracks. However, as targets are close, the tracker has *track switches* as measurements from one track and assigned to a different track.

Confusion Matrix Results for the entire run:

197	3	0
5	193	2
5	21	174

	Track 1	Track 2	Track 3
Track Purity	0.985	0.965	0.870
CAR	0.952	0.889	0.989

Note that WATP = 0.940. From these results we see that Track 2 has a high purity but lower CAR. Track exhibits the opposite analysis with a low purity and high CAR. Track 1 moves in linear fashion, while Track 2 and 3 are maneuvering. Track 2 has a short overlap with Track 1 in the beginning, while Track 3 has a large overlap period from which there is confusion with Track 2. The key here is that TP shows the incorrect assignment of measurements (*Track confusion*) while CAR demonstrates the how the target can be confused with the other targets (*ID confusion*). Since a MOE includes situational awareness, both track and ID information is required.

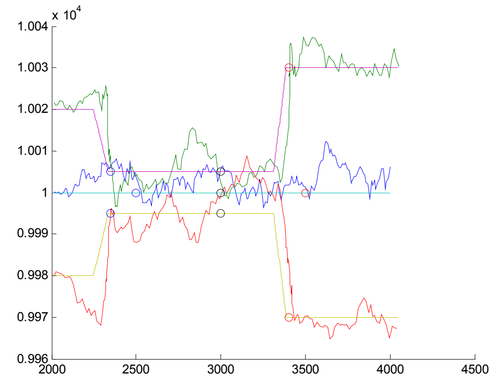

Figure 7. Tracking without ID information.

Figure 8 shows that both TP and CAR are improved with a STID system.

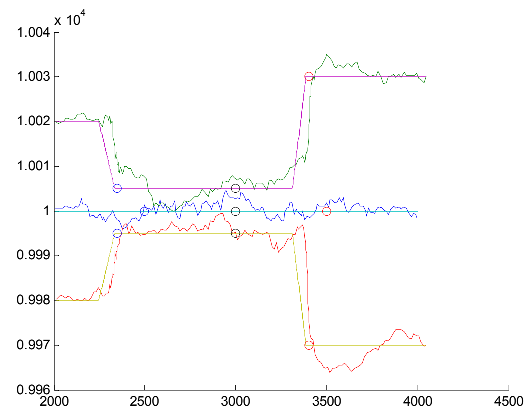

Figure 8. Tracking with ID results.

	Track 1	Track 2	Track 3
Track Purity	0.975	0.995	0.975
CAR	0.980	0.996	1.000

3.3 Measures of Performance

In an effort to do an analysis for MOEs, we plot a spider chart of the other metrics (normalized to 1 over the results for each metric), as shown in **Figure 9**. Space limits a detailed analysis; however (1) *accuracy* is high because the tracks and known, (2) *confidence* is from the ID information, (3) *timeliness* is from the measurement reporting, (4) *throughput* is from the usefulness of the data (good measurements), and (5) *value* is related to the opportunity cost. Given the area of coverage, the value of the situational analysis requires analyzing the entire space. Here it is lower due to the scenario in which the analysis of the closely spaced targets requires a focused attention of the sensors while giving up coverage over the entire area.

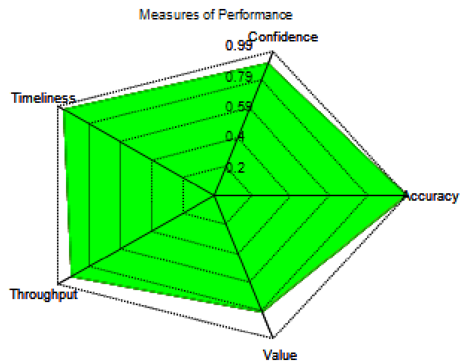

Figure 9. Measures of Performance.

MOEs are also important as related to the preference of the user to have a surveillance picture of the region of coverage, as shown in **Figure 10**.

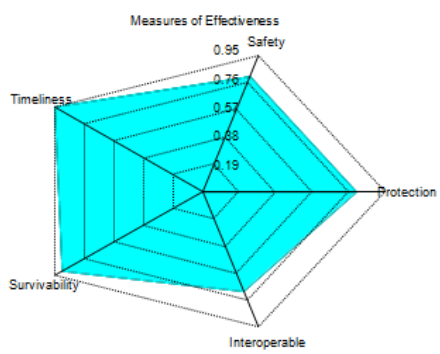

Figure 10. Measures of Effectiveness.

Timeliness is similar to the MOP, yet this is for the MOE information. *Survivability* is related to the detection of the targets, and *safety* (Safety = 1 – risk) is related to knowing all the target types and identities. Likewise, *protection* (Protection = 1 - Threat) is related to the area of coverage. *Interoperable* is related to the communication and delivery of the correct target/track information. Note that with a limited simulation, the normalized metrics highlight any discrepancies. For instance, over 200 position measurements there are numerous instances for timeliness and survivable updates. However, safety, protection, and interoperable are simulated as related to the track (versus each measurement). We will explore these metrics in future papers but are presented for discussion.

4 Discussion & Conclusions

Performance evaluation of classification results and target tracking techniques has posed difficulty in the standardization of the information. We presented numerous efforts to develop metrics and tools for evaluation. To further extend knowledge in the area, we looked at the TP metric and determined that it could be used for higher-level fusion analysis as supporting MOEs. We examined the STID scenario using the belief filter to highlight the usefulness of a *Current Assignment Ratio* in addition to a Track Purity metric.

The research aim is for system-level information fusion evaluation over the sensor, target, and environmental operating conditions and variations. We initiated a discussion on the presentation of all MOPs and MOEs in a spider plot for a user to grasp the complete performance of the STID system. Future work will explore the sensitivity of the results, the presentation of the MOE metrics, and use of operational data to validate the approach. For example, we seek to explore sensor management, image fusion, and terrain updates [**65**] as impacting the MOEs.

5 References

[1] E. Waltz and J. Llinas, "System Modeling and Performance Evaluation," Ch 11 in *Multisensor Data Fusion Systems*, Artech House 1990.

[2] Y. Bar-Shalom & X. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*, YBS, New York, 1995.

[3] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*, Artech House Publisher, Boston, 1999.

[4] J. Llinas, "Assessing the Performance of Multisensor Fusion Processes," Ch 20 in *Handbook of Multisensor Data Fusion*, (Eds.) D. Hall and J. Llinas, CRC Press, 2001.

[5] D. L. Hall and S. A. McMullen, *Mathematical Techniques in Multisensor Data Fusion*, Artech, 2004.

[6] E. Blasch, I. Kadar, J. Salerno, M. M. Kokar, S. Das, G. M. Powell, D. D. Corkill, and E. H. Ruspini, "Issues and challenges of knowledge representation and reasoning methods in situation assessment (Level 2 Fusion)", *J. of Advances in Information Fusion,* Dec. 2006.

[7] S. Mori, K. C. Chang, and C. Y. Chong, "Performance Analysis of Optimal Data Association with Applications to Multiple Target Tracking," Chapter 7 in *Multitarget-Multisensor Tracking: Applications and Advances*, Vol. II, Y. Bar-Shalom, ed., pp. 183 - 235, Artech House, 1992.

[8] X. R. Li and Y. Bar-Shalom, "Detection threshold selection for tracking performance optimization," *IEEE Tr. Aero. Elec. Sys*., Vol. AES-30, No. 3, July 1994, pp. 742–749.

[9] K.C. Chang, E. Sivaraman, and M. Liggins, "Performance Modeling for Multisensor Tracking and Classification," in *Proc. SPIE, Vol 5429*, Orlando, Florida, April, 2004.

[10]C-Y. Chong and S. Mori, "Metrics for Feature-Aided Track Association," *Fusion06*, 2006.

[11] S. Mori, K.-C. Chang, C.-Y. Chong, and K.P. Dunn, "Tracking Performance Evaluation: Prediction of Track Purity," *Proc. SPIE, Vol. 1096*, 1989.

[12]S. Mori, K. C. Chang, C. Y. Chong, and K. P. Dunn, "Prediction of Track Purity and Accuracy," *IEEE Trans. on Automatic Control*, Vol. 40, No. 5, pp. 953-959, May, 1995.

[13]C. Gokberk, "Performance Evaluation of Multisensor Data Fusion Systems," Scientific and Technical Report, May 24, 1999, Turkey.

[14] L. D. Smith, A. Register, W. D. Blair, and M. Levedahl, “A Track Purity Approach for Tracking Metrics,” *IEEE Aerospace Conf*, MT, 2010.

[15] P. Hanselman, C. Lawrence, E. Fortunano, B. Tenney, and E. Blasch, “Dynamic Tactical Targeting,” *Proc . SPIE* , Vol. 5441, April 2004.

Track and ID

[16] J. Dezert, "Autonomous navigation with uncertain reference points using the PDAF", Chapter 9 in *Multitarget-Multisensor Tracking : Applications and Advances, Vol 2*, pp 271-324, Y. Bar-Shalom Editor, Artech House, Dec. 1991.

[17] E. P. Blasch, “Flexible Vision-Based Navigation System for Unmanned Aerial Vehicles,” *Proc SPIE,* Vol. 2352, 1994.

[18]D. Salmond. D. Fisher, & N. Gordon, "Tracking and Identification for closely spaced objects in clutter," in Proc. *European Contr. Conf*., Brussels, Belgium, 1997.

[19] K. C. Chang and Robert M. Fung, “Target Identification with Bayesian Networks in a Multiple Hypothesis Tracking System”, *IEEE Aerospace and Electronic Systems*, Vol. 36, No. 3, 1997.

[20] E. P. Blasch and L. Hong “Simultaneous Identification and Track Fusion,” *Conf. on Dec. Control,* Orlando, FL, Dec 1998,

[21]E. Blasch, and L. Hong, “Data Association through Fusion of Target Track and Identification Sets,” *Fusion00*, 2000.

[22]T. Connare, E. Blasch, J. Schmitz, F. Salvatore, and F. Scarpino, “Group IMM tracking utilizing Track and Identification Fusion,” *Proc. Workshop on Est., Track, & Fusion; Trib. to YBS*, 2001.

[23]E Blasch and C Yang, “Ten methods to Fuse GMTI and HRRR Measurements for Joint Tracking and ID”, *Fusion 04*, July 2004.

[24]A. Tchamova, J. Dezert*, et al*, “Target Tracking with Generalized data association based on the general DSm rule of combination,” *Fusion04*.

[25]D. Angelova & L. Mihaylova, “Joint Tracking and Classification with Particle Filtering and Mixture Kalman Filtering using Kinematic Radar Information,” *Digital Signal Processing*, 2005.

[26]H. Chen and R. Yanhuan, “Joint target recognition and tracking using class specific features,” *IEEE Signals, Systems and Computers Conf.* 2004.

Y. Bar-Shalom and T. Kirbuarajan, and C. Gokberk, “Tracking with Classification-aided Multiframe data association,” *IEEE Trans. AES*, Vol. 43, No. 3, pp. 868-878, 2005.

[27]H. Chen, G. Chen, E. Blasch, and T. Schuck, "Robust Track Association and Fusion with Extended Feature Matching", in *Optimization & Cooperative Control. Strategies,* M.J. Hirsch et al. (Eds.):, LNCIS 381, Springer-Verlag Berlin Heidelberg 2009.

[28]E. Blasch and B. Kahler, “Multi-resolution EO/IR Tracking and Identification” *Fusion 05*, July 2005.

[29]T. Wang, Z. Zhu, and E. Blasch, “Bio-Inspired Adaptive Hyperspectral Imaging for real-Time Target Tracking “ *IEEE Sensors Journal*, Vol. 10, No. 3, pp. 647-654, 2010.

[30]J. Dezert and B. Pannetier, “A PCR-BIMM filter for maneuvering target tracking,” *Fusion10*, 2010.

[31] Z. Ding & L. Hong, “Decoupling probabilistic data association algorithm for multiplatform multisensor tracking,” *Optical Engineering*, Vol. 37, No. 2, Feb. 1998.

Fusion Evaluation

[32] E. Blasch, “Level 5 (User Refinement) issues supporting Information Fusion Management,” *Fusion06*, 2006.

[33] R.M.H. Burton, “A Survey of MOP/MOE for Maritime Command, Control and Information Systems,” *DND Operational Research and Analysis*, DMOR Research Note 9502, September 1995.

[34] E. Bossé, and J. Roy, “Measures of Performance for the ASCACT Multisensor Data Fusion, Demonstration Program,” *DREV-TM-9618*, Defence Research Establishment Valcartier, Quebec, 1997.

[35] D.S. Alberts, J. J. Garstka, and F.P. Stein, “Network Centric Warfare, Developing and Leveraging Information Superiority,” *DoD C4ISR Cooperative Res. Pro., CCRP pub. series*, 1999.

[36] A. Theil, L.J.H.M. Kester and E. Bossé, “On Measures of Performance to Assess Sensor Fusion Effectiveness,” *Fusion 2000*, 2000.

[37] Y. Tessier, P. Valin, D. Turgeon, V. Auns, R. Hallsworth, “Measures of Performance and Measures of Effectiveness for COMDAT 1”, LM Report, 6520003004, 2001.

[38]E. Blasch, M. Pribilski, B. Roscoe, *et. al.*, “Fusion Metrics for Dynamic Situation Analysis,” *Proc SPIE 5429*, April 2004.

[39]I. Kadar, K. C. Chang, K. O’Connor, and M. Liggins, “Figures-of-Merit to bridge Fusion, long-term prediction and dynamic sensor management,” *Proc SPIE 5429*, April 2004.

[40] K. C. Chang, “Analytical Performance Evaluation for Autonomous Sensor Fusion,” in *Proc. SPIE*, March, 2008.

Track Performance Evaluation

[41]Y. Bar-Shalom, K. C. Chang, and H. M. Shertukde, “Performance Evaluation of a Cascaded Logic for Track Formation in Clutter,” *IEEE Trans. on Aerospace and Electronic Systems*, Nov., 1989.

[42]K. C. Chang, S. Mori, and C-Y. Chong, “Performance Evaluation of Track Initiation in Dense Target Environments,” *IEEE Trans. AES*, Vol. 30, No.1, 1994.

[43]K. C. Chang, S. Mori, and C-Y. Chong, “Evaluating a Multiple-Hypothesis Tracking Algorithm,” *IEEE Trans. AES*, Vol. 30, No.2, 1994.

[44] K. C. Chang, R. Saha, and Y. Bar-Shalom, “On Optimal Track-to-Track Fusion,” *IEEE Trans. on Aerospace and Electronics Systems*, Vol. 33, No. 4, pp. 1271-1276, Oct., 1997.

[45] R.L. Rothrock and O.E. Drummond, “Performance Metrics for Multiple-Sensor, Multiple-Target Tracking,” *Proc. of SPIE Vol. 4048*, 2000.

[46] X. R. Li, Z. Zhao, and V. P. Jilkov. Practical Measures and Test for Credibility of an Estimator. In *Proc. Workshop on Estimation, Tracking, and Fusion: A Tribute to Yaakov Bar-Shalom*, pages 481–495, Monterey, CA, May 2001.

[47]K. C. Chang, Z. Tian, and R. Saha, “Performance Evaluation of Track Fusion with Information Matrix Filter,” *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 38, No. 2, April, 2002.

[48]K. C. Chang, Z. Tian, S. Mori, and C-Y Chong, “MAP Track Fusion Performance Evaluation,” Fusion02, 2002.

[49]K. C. Chang, Z. Tian, S. Mori, and C-Y. Chong, “Performance Evaluation for MAP Track Fusion,” *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 40, No. 2, pp. 706-714, April, 2004.

[50] X. R. Li and Z.-L. Zhao. Evaluation of Estimation Algorithms—Part I: Incomprehensive Performance Measures. *IEEE Trans. Aerospace and Elec. Sys.*, AES-42(4):1340–1358, Oct. 2006.

[51] X. R. Li and Z.-L. Zhao. Testing Estimator’s Credibility—Part I: Tests for MSE. In *Proc. Intl Conf. on Info. Fusion*, 2006.

[52] X. R. Li and Z.-L. Zhao. Testing Estimator’s Credibility—Part II: Other Tests. In *Proc. Intl Conf. on Info. Fusion*, 2006.

[53] G. Wah Ng, C. Huat Tan, T. Poh Ng, and S. Yi Siow, “Assessment of Data Fusion Systems,” *Fusion06*, 2006.

[54] O. Straka, M Flidr, J. Dunik, M, Simandl, and E. P. Blasch, “Nonlinear Estimation Framework in Target Tracking,” *Fusion10*, 2010.

[55] K. C. Chang, C-Y. Chong, and S. Mori, “Analytical and Computational Evaluation of Scalable Distributed Fusion Algorithms,” *IEEE Trans. on AES*, Vol. 46, No. 4, Oct., 2010.

Classification Fusion Evaluation

[56] S. Alsing, E. P. Blasch, and R. Bauer, “3D ROC Surface concepts for Evaluation of Target Recognition algorithms faced with the Unknown target detection problem,” *SPIE Conf. 3718,* 1999.

[57] E. Sivaraman and K. C. Chang, “The Global Confusion Matrix: A New Approach to Evaluating the Performance of Multi-sensor Classification Systems”, *Proc. Fusion 2005*, July, 2005.

[58] E. Sivaraman and K. C. Chang, “Performance Evaluation of Multi-Sensor Classification Systems,” *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 43, pp. 1265-1281, Oct., 2007.

[59] R. S. Lynch and P. K. Willett, “Bayesian Classification and Feature Reduction Using Uniform Dirichlet Priors,” *IEEE Trans. Sys., Man, and Cybernetics, Part B*, Vol. 33, No. 3, June 2003.

[60] P. Djiknavorian, D. Grenier and P. Valin, “Analysis of information fusion combining rules under the DSm theory using ESM inputs”, *Fusion07*.

[61] R. N. Carvalho and K. C. Chang, “A performance evaluation Tool for Multi-Sensor Classification Systems,” *Fusion09*, 2009.

SA Performance
[62]E. P. Blasch "Assembling an Information-fused Human-Computer Cognitive Decision Making Tool," *IEEE AESS Mag.,* 2000.
[63]J. Salerno, E. Blasch, M. Hinman, and D. Boulware, "Evaluating algorithmic techniques in supporting situation awareness," *SPIE 05*, 5813, April 2005.
[64] E. P. Blasch, P. Valin, and E. Bossé, "Measures of Effectiveness for High-Level Fusion," *Fusion10*, 2010.
[65] *NATO Code for Best Practice (COBP) for C2 Assessment*, 2002. A DoD CCRP/NATO Collaboration
[66] C. Wallshein, NATO Code for Best Practice (COBP) for C2 Assessment, 2002. (Slides)
[67] S. H. Star, NATO Code for Best Practice (COBP) for C2 Assessment, 2002. (Slides)

[68] C. Yang and E. Blasch, "Kalman Filtering with Nonlinear State Constraints," *IEEE Transactions AES*, Vol. 45, No. 1, Jan. 2009.