Performance Assessment of Sonar-System Networks for Anti-Submarine Warfare

J.M. Thredgold, M.P. Fewell, S.J. Lourey and H.X. Vu
Defence Science and Technology Organisation, Maritime Operations Division
P.O. Box 1500
Edinburgh SA 5111, Australia

***Abstract*- This paper describes a study of the effect of networking active sonars on tracking performance in a scenario with explicit inclusion of false detections, using two false-detection models and detection-probability curves with two levels of tailing. We compare the tracking performance when sonars share detections (centralized tracking) with the performance when sonars share tracks (distributed tracking). Provided that the sonar layout and detection probabilities are such that multiple sonars have a reasonable probability of obtaining detections from a target, we show that centralized tracking decreases the time to confirm a track on a target and improves the continuity of the target track, but increases the false track rate. Results for other metrics are also presented.**

I. INTRODUCTION

The prevailing paradigm governing concepts of command and control—network-centric warfare [1,2]—leads one to ask whether networking can assist in antisubmarine-warfare (ASW). Suppose a sonar network has the optimum number of nodes optimally connected with all the bandwidth desired, plus any other optimistic assumption thought useful. Is it better, in terms of military operational effectiveness, to share information on tracks ("distributed tracking") or on detections ("centralized tracking")? This paper presents part of an ongoing quantitative study of this question.

The question pertains to multisensor data fusion, which has been much studied in the above-water domain (for reviews, see [3,4]). It has long been known that pooling detections gives better performance in theory [5], but no general conclusion is possible if the contributing detectors have differing detection probabilities and false-detection rates [6]. Data fusion can make things worse [3], as demonstrated by some long-standing counterexamples [7,8]. The authors of a recent large study commented that a completely distributed approach is not feasible but a completely centralized approach is neither robust nor scalable [9]. In short, it seems necessary to examine each case on its merits; here, we focus on the requirements of ASW by active sonar.

Previously we had sought to quantify operational gain (or loss) in pooling detections in terms of coverage area enclosed by contours of constant track-initiation probability [10]. That study shows the importance of the shape of the single-ping detection probability (p_d) – range curve: networking advantage is high if the curve decreases slowly with range, but goes to zero as the curve approaches a "cookie-cutter." In separate, though related work, we examined the utility of defining detection range in terms of cumulative track-initiation probability [11]. Apart from the manifest neglect of later stages of the tracking process, both studies avoided the issue of data association by, among other things, not explicitly including false detections in the modeling. The present study is a first step toward addressing these issues.

The previous work also shows that, when advantage in centralizing tracking is obtained, it can be proportionally largest in very small networks [10]. Hence, we consider just three sonar systems here. We limit the networking to detection-level at most; that is, the sonars operate multiple-monostatically. Consideration of multistatic systems is left for future study.

We expect that pooling detection information will always lead to an elevated false-track rate. Unfortunately, the accurate modeling of false detections in sonar systems is an unsolved problem (e.g. [12]). Here, we try two models. The first is the usual model comprising a Poisson-distributed number of detections per ping randomly and uniformly distributed over the field of view. For the second, we add ping-to-ping persistence by randomly selecting a nominated fraction of false detections to retain for the following ping.

II. METHOD

Detection data are generated by Monte-Carlo iterations of a scenario comprising 3 sonar systems that encounter a submarine positioned a little off their line of advance. The initial position of the platforms is shown in Fig. 1. The sonars maintain a regular ping sequence, each pinging once per minute. At each ping, the p_d value for the range concerned is used to determine whether a detection occurred. Each detection generates a record of the submarine's position modified randomly to represent measurement errors. Each ping also generates false detections randomly and uniformly distributed over the field of view. The number of false detections at each ping is determined by a draw from a Poisson distribution with a mean corresponding to a false-alarm rate of 10^{-5} per range–bearing cell per ping. When ping-to-ping persistence is included, the specified percentage of false detections from the previous ping of a sonar is randomly selected and retained, with their locations jittered, and new false detections are generated, if necessary, to make up the number required for the new ping.

978-1-4244-5222-4/10/$26.00 ©2010 IEEE

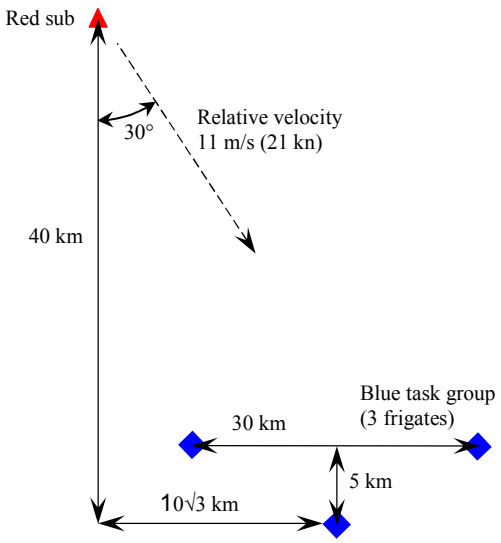

Figure 1. The positions of the three sensors and the submarine at the beginning of the simulation. The relative velocity is constant throughout the simulation.

The result of one iteration of the simulation is a list of detections, both target and false, for each sonar system. These are then run through the tracker, either individually (distributed tracking) or after fusing to form a single list (centralized tracking). The tracker is a simple Kalman filter with a 3-in-5 rule for both track initiation and termination. Data association is by nearest neighbour in Mahalanobis distance from the predicted position, using a 95% confidence level that the detection is related to an active track. All unassociated detections are candidates for forming new tracks. In the distributed tracking case, the three track lists are then fused, with coincident tracks being converted to a single track. Finally, since a track can include both target and false detections, tracks containing more target than false detections are classified as "target", otherwise they are "false". Full details of the simulation and tracker are available elsewhere [13].

A set of metrics is then applied to the two lists of tracks—distributed and centralized—to quantify operational performance. The metrics used are:

- target-track establishment delay,
- number of false tracks,
- length of both false and target tracks,
- total number of tracks at any given time, and
- system confusion matrix, using a classification rule based on track length.

Details of the metrics are given in [14]. Each metric is computed separately for the two track lists, with statistics accumulated over 1000 Monte-Carlo iterations.

The above process is carried out for each of the following trial conditions:

- a "Fermi" p_d curve with a relatively sharp cut-off, approaching a cookie-cutter shape,
- an exponential p_d curve, which has a long tail out to relative large range, as in [14], and
- the exponential p_d curve with persistent false detections

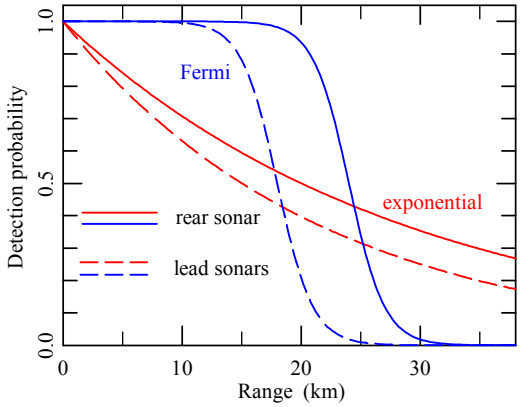

Figure 2. The detection-probability curves used.

at six levels in the range 5%–50%.

The p_d curves used are shown in Fig. 2. For the purposes of the study, the rear sonar is given a longer detection range.

III. Results

A. *Track Establishment Delay*

Track establishment delay is defined as the number of pings between the first detection of the target and the initia-

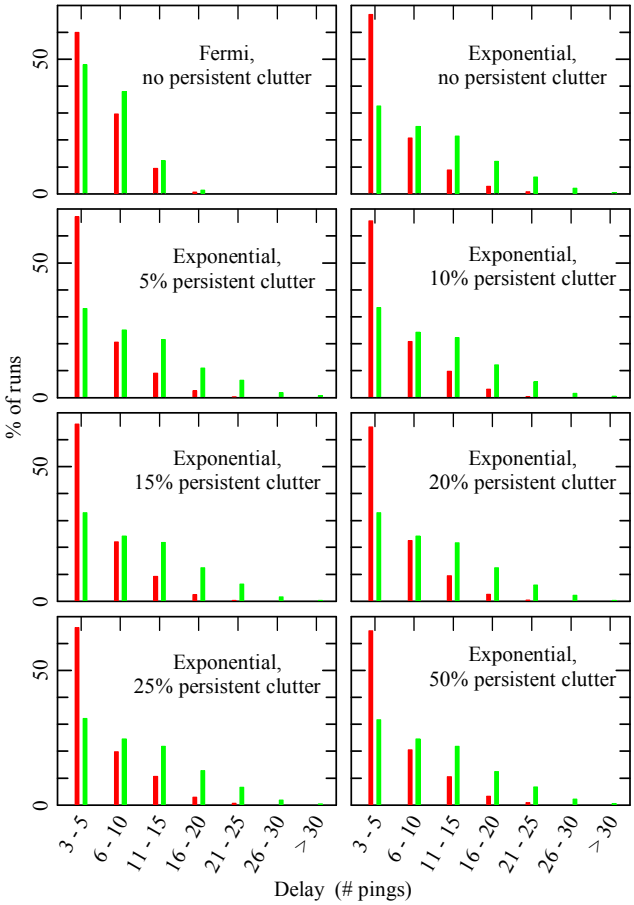

Figure 3. Distributions of track-establishment delay for the eight trial conditions, each compiled from 1000 Monte-Carlo iterations; red bars—centralized tracking; green bars—distributed tracking

tion of a track. Fig. 3 shows histograms of this quantity for the eight trial conditions studied. Since we use a 3-in-5 track-initiation rule, the delay must be at least 3 pings, so the first bin of each distribution starts at 3 pings. It is clear that centralized tracking begins earlier on average for all trial conditions, with fewer Monte-Carlo iterations where the delay is large, but it is difficult to discern how the behavior varies with the trial conditions. For this, we examine distributions of differences between centralized and distributed tracking. That is, for each Monte-Carlo iteration, we compute the difference in track-establishment delay between the two tracking modes. Fig. 4 shows examples of distributions of these differences. In every trial condition, including those not shown in Fig. 4, the most common case is zero delay difference; for these iterations tracking begins on the same ping for both tracking modes. However, for the 60–75% of the iterations where there is a difference, it is usually the centralized tracking that starts earlier. That is, although each trial condition has iterations in which the distributed-mode tracking starts before the centralized tracker, predominantly it is the other way around.

Fig. 5 summarizes the results. (Note that the thick bars in Fig. 5 indicate the 5–95 percentile range rather than the first–third quartile range more common in "whisker plots.") For the exponential p_d curve, centralized tracking starts 4–5 pings earlier, on average, than distributed tracking. Fig. 5 shows the extent to which track-initiation delay difference is insensitive to persistence in the false detections. On the other hand, the Fermi (cookie-cutter like) p_d curve leads to less advantage for centralized tracking — initiation is earlier by a mean of just 1 ping — than the exponential curve. This is also evident in Fig. 4.

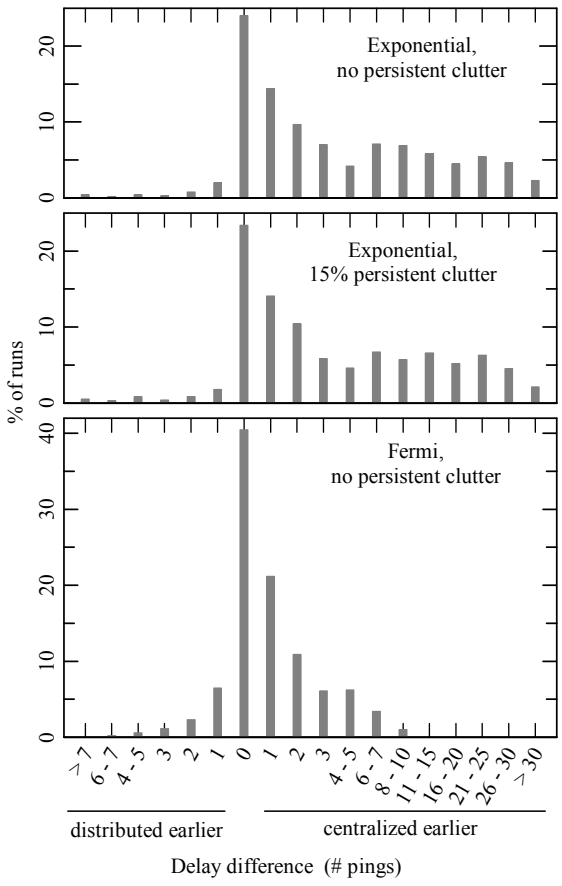

Figure 4. Distributions of track-establishment-delay difference for three of the eight trial conditions.

B. *False Tracks*

The metrics recorded for false tracks were the number of false tracks occurring in a simulation iteration and the lengths of these false tracks. Fig. 6 shows the distributions of the number of false tracks generated by both centralized and distributed tracking for the eight trial conditions. As expected more false tracks — about 4 times more for all trial conditions — are formed by centralized tracking, and the number of false tracks also increases as the persistence of the false detections is increased. The increase in false track numbers

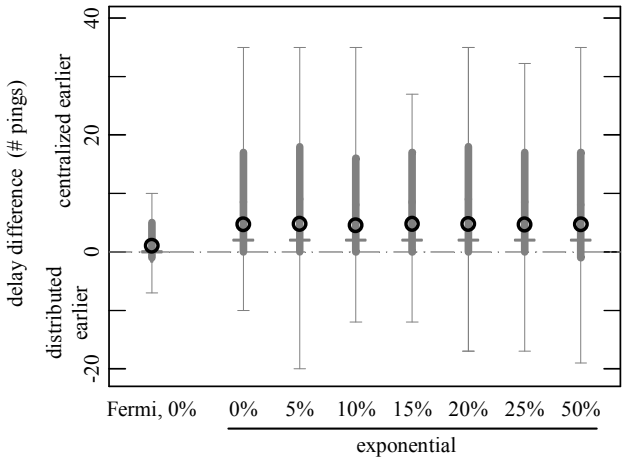

Figure 5. Summary statistics of distributions of track-establishment delay difference for the eight trial conditions; whiskers—min–max range; thick bars—5th–95th percentile range; horizontal marks—medians; black circles—means

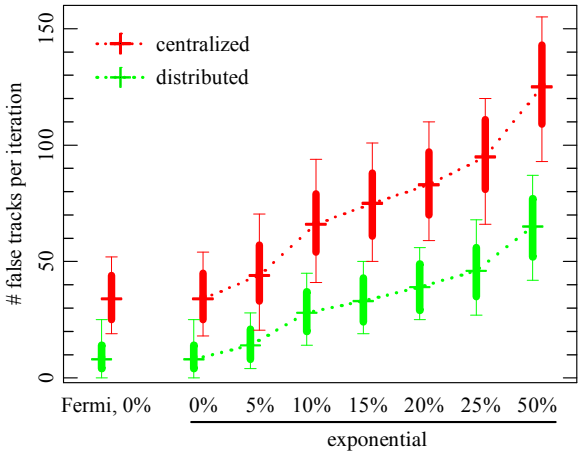

Figure 6. Summary statistics of distributions of number of false tracks for the eight trial conditions; whiskers—min–max range; thick bars—5th–95th percentile range; horizontal marks—medians

with false-detection persistence appears linear for both centralized and distributed tracking. (Linear regression gives correlation coefficients—R^2 values—of 0.95 for both the centralized and distributed cases. Note that the inclusion of the 50% persistence values means that the scale on the x-axis of Fig. 6 is not linear.) The Fermi (cookie-cutter like) probability of detection curve shows an increase in the number of false detections for centralized tracking similar to that for the exponential curve with no persistent false detections.

Unlike the number of false tracks, false track length does not appear to vary with either the persistence of false detections or between the Fermi and exponential p_d curves. Fig. 7 shows that the maximum false track length increases with increasing proportion of persistent false detections, but these are outliers occurring in less than 5% of the iterations; the majority of false tracks, even with 50% persistence of false detections, are 6 or fewer pings in length. Close inspection of the distributions of false track length fails to reveal any discernable difference between centralized and distributed tracking for any trial condition.

C. *Target-track Length*

As for false tracks, we recorded the length of each target track formed. However rather than present the distribution of target track lengths, Fig. 8 shows the percentage of target tracks that exceed 15 and 30 pings in length. The much higher number of long target tracks formed by the central tracker indicates that it is easier to maintain a track when tracking is centralized. Also centralized tracking does not appear to be adversely affected by increasing persistence of false detections. Distributed tracking, in contrast, shows a decrease in the number of long target tracks with increasing persistence of false detections.

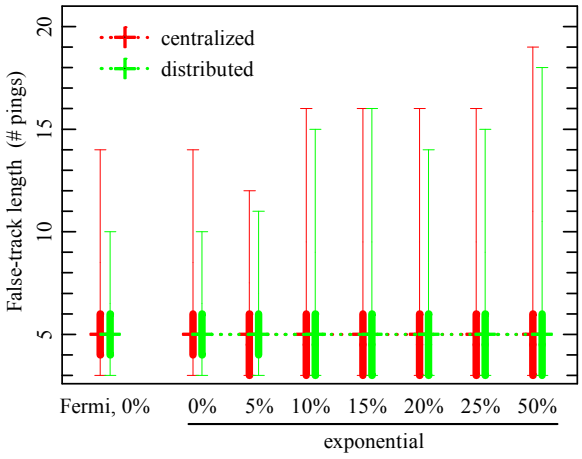

Figure 7. As in Fig. 6, but for false track length

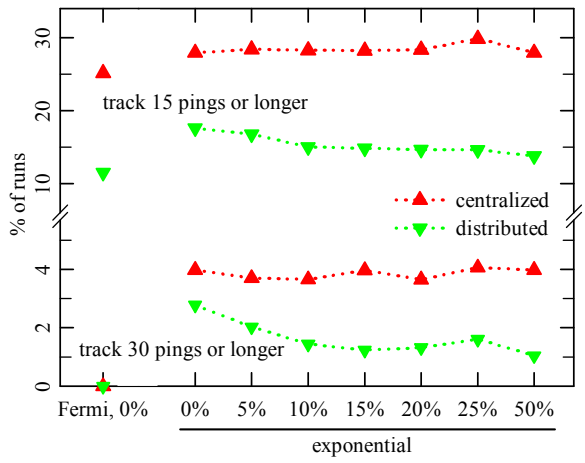

Figure 8. Length of the target track for the eight trial conditions

For the Fermi-function probability of detection, there are no target tracks which reach 30 pings in length. This is consistent with the delay in track initiation; for the Fermi function case tracks do not generally start early enough for them to be maintained for 30 pings during the simulation, which has a total duration of 42 pings.

D. *Instantaneous Total Number of Tracks*

The main reason for concern regarding the number of false tracks is that these have a significant impact on the workload of the sensor operator. While the total number of false tracks is an important indicator of this, the number of tracks in existence at any given time is also a measure of operator workload at that time. Fig. 9 shows summary statistics for the total number of tracks in existence at each ping during the simulation for one trial condition (exponential with no persistent clutter) and centralized tracking. Start and end effects are apparent for about 5 pings from the beginning and end of the simulation.

In the example shown in Fig. 9, for most of the simulation run time there is a mean number of about 7 tracks active, with 90% of the iterations having between 3 and 11 tracks active at any given time mid iteration. Fig. 10 shows that these numbers all increase as the false-detection persistence increases. The mean number of tracks is as many as 22 for centralized tracking with 50% persistence of false detections.

The total number of tracks is consistently higher for centra-

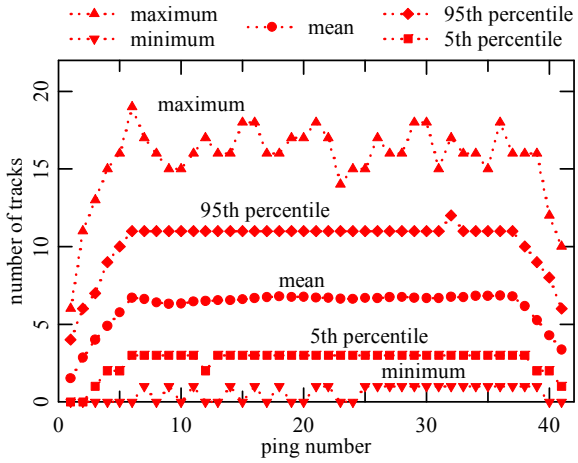

Figure 9. Summary statistics of distributions of the number of tracks, both false and target, at each ping during an iteration of the scenario with the exponential p_d curve, no persistent clutter and centralized tracking

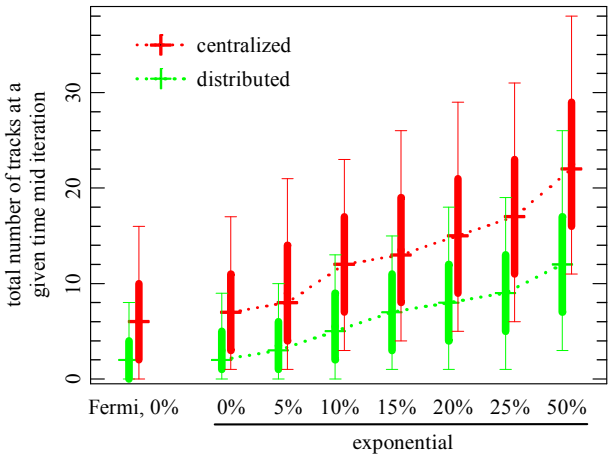

Figure 10. Estimates of the summary statistics of instantaneous track numbers mid iteration for the eight trial conditions

lized tracking than it is for distributed, regardless of the probability of detection curve or the persistence of false detections. With zero persistence of false detections, centralized tracking leads to about 2.5 times more tracks active at any given time than distributed tracking for both p_d curves. This proportion falls to a little under 2 at 50% persistent false detections. This increased false track rate arising from networking may cause the operator to adjust the detection threshold to reduce the false track rate to a more manageable level, but thereby reducing the detection probability actually achieved.

E. System Confusion Matrix

The system confusion matrix provides a measure of the proportion of target and false tracks that would be classified correctly considering track length alone. In reality this would not be the sole classification technique used, but it may be used by an operator to prioritize or filter the tracks investigated, especially if there are a large number of tracks. Fig. 11 shows the number of tracks wrongly classified with the classification rule: all tracks of 5 or fewer pings in length are false; all tracks maintained for at least 6 pings are target tracks.

As expected, Fig. 11 shows that track length becomes less useful as a classification tool for false tracks as the level of persistent false detections increases (i.e., the percentage of misclassified false tracks increases for both tracking modes). There is a corresponding increase, at least initially, in the number of target tracks which are wrongly classified for distributed tracking, but we do not see the same increase for centralized tracking; in fact the reverse: as false-detection persistence increases, the effectiveness of track length as a classification rule for target tracks improves. This results from the concentration of false detections. As the persistence increases, false detections tend to "cluster" and are no longer uniformly spread across the sonar's field of view. This reduces the likelihood of false detections interfering with the formation or maintenance of target tracks. There may be a number of reasons why similar behavior is not seen for the distributed tracking. Perhaps the clustering has less effect because the total number of false detections is lower.

For the Fermi p_d curve, target track classification by length is much more accurate (for both centralized and distributed tracking) than it is for the exponential curve (i.e. the proportion of misclassified target tracks is much lower — left-most cluster of points in Fig. 11). This is due to the cookie-cutter-like nature of the Fermi curve. With the exponential p_d curve there may be a number of short, stop-start target tracks in the "tail" region. For the Fermi curve, it generally takes longer to begin a track, but once a track is formed it is much more likely to be continuously maintained as the probability of detection is high. This highlights the need for accurate classification techniques in order to take advantage of the earlier tracking opportunities that may be available when networking.

Despite the change in the effectiveness of the track length as a classification tool as false-detection persistence increases,

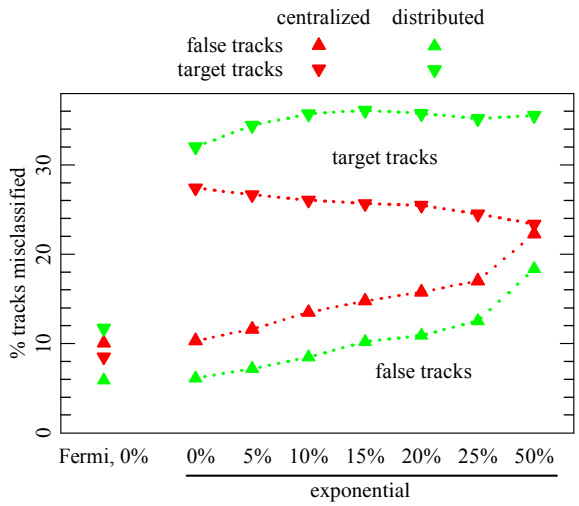

Figure 11. Values of two elements of the confusion matrix: percentages of target tracks and false tracks misclassified when the classification rule is based on a track length of 5 pings.

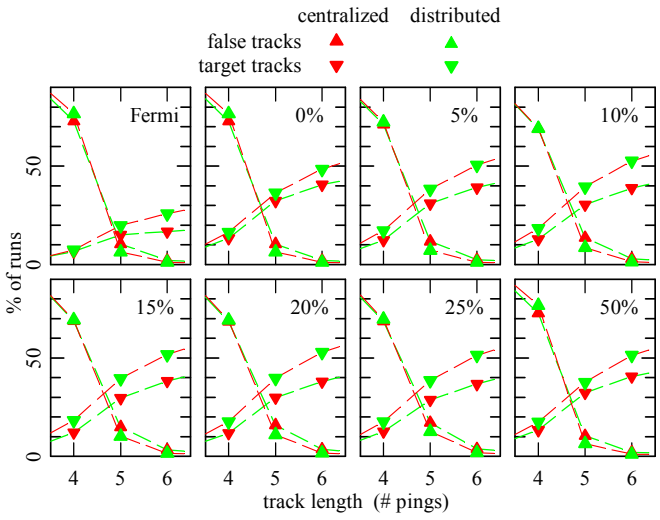

Figure 12. Basis for selecting 5 pings as the track length for the track-classification rule. The optimal length for classification is the integer value closest to the intersection of the relevant lines.

Fig. 12 confirms that 5 pings is the best track length on which to classify for all trial conditions.

IV. Summary

The purpose of this study is to compare the performance of fully centralized and fully distributed tracking modes in an ASW scenario. The various trial conditions explore this comparison for two different p_d curves — with a substantial long-range tail and without — and also with and without persistent false detections. This provides the following three dimensions for systematizing the results.

A. Centralized versus Distributed Tracking

Focusing on the results for the exponential p_d curve without persistent false detections leads to the following conclusions on the benefits or otherwise of centralizing the tracking process:

- Centralized tracking produces an earlier start to the tracking, by an average 4–5 pings.
- On the other hand, centralized tracking produces about 4 times the number of false tracks as distributed tracking, for the scenario parameters chosen here. The total number of tracks active at any time during a scenario run is about 2.5 times higher for centralized than distributed tracking.
- Both tracking modes give false tracks of the same length, on average.
- Centralized tracking leads to longer target tracks or, in other words, to less fragmentation of the target track.
- If track length is used as the sole basis for classifying tracks, then a length of 5 gives the best results. Centralized tracking misclassifies fewer target tracks but more false tracks (i.e. centralized tracking results in fewer missed targets but more false alarms).

B. Effect of Tailing in the p_d Curve

Comparison of the results with exponential and Fermi p_d curves shows their sensitivity to the level of tailing in the p_d curve. We see that:

- The extent to which centralized tracking produces earlier track initiation is much reduced, from an average of 4–5 pings with an exponential p_d curve to just over 1 ping with the Fermi curve.
- On the other hand, the relativities of the number of false tracks is unaffected by p_d-curve tailing: centralized tracking produces about 4 times the number of false tracks as distributed tracking with either p_d curve, and about 2.5 times the number of tracks active at any given time.
- Both tracking modes continue to give false tracks of the same length, on average.
- Centralized tracking also continues to give longer target tracks than distributed tracking, but the fraction of long tracks is much lower for the Fermi p_d curve than the exponential.
- As regards using track length for classification, this works much better for the Fermi p_d curve than the exponential: all misclassification rates are reduced, those for target tracks by about a factor of 3, rather less for false tracks. The relativities between centralized and distributed tracking are approximately maintained: centralized tracking still misclassifies fewer target tracks but more false tracks. A track length of 5 still gives the best results.

C. Effect of Persistent False Detections

Now we turn to the results as the level of persistent false detections is increased (exponential p_d curve only).

- There is no discernable effect on the extent to which centralized tracking produces earlier track initiation. It remains at an average of 4–5 pings even with half of the false detections being persistent.
- The total number of false tracks in a scenario rises linearly with the percentage of persistent false detections, but the relativities between exponential distributed tracking are little affected: at the highest percentage, centralized tracking produces about twice the number of false tracks as distributed tracking, compared with 4 times when none of the false detections is persistent.
- False-track length is unaffected by persistent false detections, and the effect on the proportion of long target tracks is small.
- The number of tracks active at any given time rises steadily with an increasing level of persistent false detections. However, we expect that the persistence of these false tracks would also be greater, so this might not represent the increase in operator workload that it may appear at first sight.
- The effect of persistent false detections on classification by track length is complex. The percentage of misclassified false tracks rises with the percentage of persistent detections for both tracking modes, at about the same absolute rate. On the other hand, the performance of centralised tracking in the classification of target tracks *improves* (i.e. the percentage of misclassified target tracks falls) with increasing persistent false detections, whereas the performance of distributed tracking becomes worse, although only slightly.

V. Conclusion

Taken as a whole, the metrics indicate that centralized tracking has advantages over distributed tracking for anti-submarine warfare with active multiple-monostatic sonar, particularly for small networks. However the use of centralized tracking with larger networks or higher false detection rates than we have considered in this example will result in a significant increase in the false track rate.

The benefits of centralized tracking are more pronounced when the probability of detection curves have long tails rather than sharp cookie-cutter like cut-offs. Persistence of false detections does not reduce the advantages of centralized tracking (earlier track initiation, greater continuity in target

tracks) but may worsen the disadvantages, causing an increase in the false track rate.

In order to fully realize the benefits of centralized tracking with a network of sonars, we will need to develop techniques and tools to help sonar operators to deal with the increased false alarm rate. These may include filtering on other sonar-signal properties (e.g. Doppler shift, signal level), fusion with other information, such as a tactical picture, or the active classification methods currently under development at DSTO [15, 16].

REFERENCES

[1] J. Bell, J. Clothier, and J. O'Neill, J "Options for Australian defence in network centric warfare," report DSTO-CR-0339 of the Defence Science and Technology Organisation, April 2003.

[2] M.P Fewell and M.G Hazen, "Network-centric warfare: its nature and modelling," report DSTO-RR-0262 of the Defence Science and Technology Organisation, September 2003.

[3] D.L. Hall and J. Llinas, "An introduction to multisensor data fusion," *Proc. IEEE*, vol. 85, pp. 6–23, January 1997.

[4] D. Smith and S. Singh, "Approaches to multisensor data fusion in target tracking: a survey," IEEE Trans. Knowledge Data Eng., vol. 18, pp. 1696–1710, December 2006.

[5] Z. Chair and P.K. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. Aerospace Elect. Sys.*, vol. AES-22, pp. 98–101, January 1986.

[6] S.C.A. Thomopoulos, R. Viswanathan, and D.C. Bougoulias, "Optimal decision fusion in multiple sensor systems," *IEEE Trans. Aerospace Elect. Sys.*, vol. AES-23, pp. 644–652, September 1987.

[7] R. Viswanathan and V. Aalo, "On counting rules in distributed detection," *IEEE Trans. Acoustics, Speech Sig. Process.*, vol. 57, pp. 772–775, May 1989.

[8] M. Cherikh and P.B. Kantor, "Counterexamples in distributed detection," *IEEE Trans. Inform. Theory*, vol. 38, pp. 162–165, January 1992.

[9] S. Oh, L. Schenato, P. Chen, and S, Sastry, "Tracking and coordination of multiple agents using sensor networks: system design, algorithms and experiments," *Proc. IEEE*, vol. 95, pp. 234–254, January 2007.

[10] M.P. Fewell, J.M Thredgold, and D.J. Kershaw, "Benefits of sharing detections for networked track initiation in anti-submarine warfare," report DSTO-TR-2086 of the Defence Science and Technology Organisation, January 2008.

[11] M.P. Fewell and J.M Thredgold, "Cumulative track-initiation probability as a basis for assessing sonar-system performance in anti-submarine warfare," report DSTO-TN-0932 of the Defence Science and Technology Organisation, December 2009.

[12] D. Grimmett and S. Coraluppi, "Contact-level multistatic sonar data simulator for tracker performance assessment," *Proc. 9th Internat. Conf. on Inform. Fusion*, 2006.

[13] J.M. Thredgold, S.J. Lourey, H.X. Vu, and M.P. Fewell, "A simulation model of networked tracking in anti-submarine warfare," report DSTO-TR-2372 of the Defence Science and Technology Organisation, January 2010.

[14] J.M Thredgold and M.P. Fewell, "Distributed versus centralised tracking in networked anti-submarine warfare," report DSTO-TR-2373 of the Defence Science and Technology Organisation,.January 2010.

[15] A. von Trojan and A. Kouzoubov, "Active sonar classification research tool capabilities and implementation outcomes', report DSTO-TR-2042 of the Defence Science and Technology Organisation, August 2007.

[16] A. von Trojan and A. Kouzoubov, 'Active sonar classification using a binary kernel machine classifier', report DSTO-TR-2110 of the Defence Science and Technology Organisation, March 2008.