



A Comparison of Global Nearest Neighbor and Linear Assignment Problem Solution Methods on a Model Tracking Problem

Roy Danchick

2051 S. Bentley Ave, Ste 302, Los Angeles, CA 90025, USA

Email: dodeee@sbcglobal.net

How to cite this paper: Danchick, R. (2024) A Comparison of Global Nearest Neighbor and Linear Assignment Problem Solution Methods on a Model Tracking Problem. *Open Access Library Journal*, 11: e12173. <https://doi.org/10.4236/oalib.1112173>

Received: August 27, 2024

Accepted: October 8, 2024

Published: October 11, 2024

Copyright © 2024 by author(s) and Open Access Library Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In this paper we compare track data association purity, accuracy, and timing on a simple, idealized model tracking problem for two data association methods: Global Nearest Neighbor (GNN) and Linear Assignment Problem (LAP). Accurate data association is the central problem of multi-target track assembly. Though simple, the model, a 1d process noise-free Kalman filter, captures the essence of the problem of tracking multiple closely spaced objects: 1) assembly of object sensor measurements into tracks in the space of measurements 2) estimation of the Kalman filter state parameters giving rise to each measurement. We show that a Jonker-Volgenant (JV) LAP method decisively outperforms GNN in all three performance measures. Moreover, our results clearly show that the use of GNN methods for data association is highly problematic. Our basic recommendation is that a Multiple Hypothesis Tracking (MHT) method, which exploits a rectangular matrix extension of the JV algorithm as the core solver of a Murty's ranked assignment algorithm, should be the preferred method for tracking multiple closely spaced objects.

Subject Areas

Applied Statistical Mathematics

Keywords

Multi-Target, Target Tracking Data, Association Global Nearest Neighbor Linear Assignment Problem, Jonker-Volgenant Algorithm, Kalman Filter

1. Introduction

Based on two disquieting experiences with the application of GNN for ballistic

missile tracking data association it comes as a surprise to us that there is still active research on this method.

The first of these experiences occurred at an early 1980's Denver conference on satellite infrared sensor ballistic missile tracking. One of the conference papers reported a GNN track-to-observation data association accuracy of 22 % in a high fidelity ballistic missile simulation.

In the late 1990's, we had already written our Multiple Hypotheses Tracking paper [1] and were readying it for submission for publication. In it we introduced the notion of track purity as a basic data association performance metric; it is the maximum of the fractions of the number of observations in an assembled track of observations all of which come from the same sensor-detected object. In that paper we reported that track data association based on a Murty [2] K-best hypotheses algorithm with a Jonker-Volgenant (JV) [3] linear assignment problem solution algorithm as its core solution method was producing track purities in the high 90 percentiles for 200 closely spaced objects over 37 scans in a satellite infrared sensor high fidelity simulation.

In a Government funded Research and Development contract at our company we extended our satellite sensor 2d line-of-sight observation data association algorithm to 3d radar data association and were testing its performance in a radar system simulation. The systems engineer, working for another company hosting the radar simulator, had instrumented track purity performance evaluation with a GNN data association solution, scan-by-scan, as ground truth. Not unexpectedly, our extended algorithm's track purity results came in at the low 20 percentiles.

So, it comes as a disturbing surprise to us that there is a still significant recent research devoted to GNN algorithms applied to tracking systems. Nowhere in this literature (e.g., references [4]-[9]) is track data association purity or an analogous metric ever mentioned, much less used to evaluate GNN data association performance.

The intent of this paper is to demonstrate, with the simplest of models, the striking, absolute performance dominance of the LAP algorithm over the GNN algorithm in terms of track data association purity, track state estimation accuracy, and timing.

Section 2 below covers definitions and notation. Section 2 provides derivations.

Section 3 supplies our key results.

Section 4 discusses the major conclusions that flow from our results.

Section 5 lists references cited in the paper.

2. Definitions

n_s = number of scans.

n_t = number of tracks.

n_o = number of observations per scan.

σ = observation normal error standard deviation.

$\boldsymbol{\mu} = [0, 1, 2, \dots, n_t - 1]^T$ is the column vector of state expected values.

x_{is} = i^{th} track state on scan s at the beginning scan s prior to state update

$i = 1, 2, \dots, n_t$.

y_{js} = j^{th} observation on the s^{th} scan., $j = 1, 2, \dots, n_o$.

$\mathbf{I} = [1, 2, 3, \dots, n_t]^T$.

$\mathbf{L} = n_t \times n_o$ track-to-observation data assignment matrix.

$\mathbf{C} = [c_2, c_3, \dots, c_{n_s}]^T$ is a column vector of costs for either the GNN or LAP methods.

\mathbf{J}_s = the assignment solution permutation column vector for either GNN or LAP methods on the s^{th} scan.

$\mathbf{P} = n_t \times n_s$ track-to-object assignment purity matrix for either GNN or LAP methods.

3. Derivations

3.1. Preliminaries

We first note that the core of our simulation model is a one dimensional Kalman filter with unit transition and measurement matrices, process noise-free, and independent identically distributed normal zero $N(0, \sigma^2)$ measurement error for each observation in each scan. The track-to-observation assignment matrix is square to facilitate the exploitation of the MATLAB JV linear assignment problem solver.

3.2. Initialization

In our simulation we first create the common $n_t \times n_s$ matrix of observations, \mathbf{Y} , with (i, s) element given by

$$y_{is} = \mu_i + e_{is}, \quad e_{is} \in N(0, \sigma^2), \quad i = 1, 2, \dots, n_t; \quad s = 1, 2, \dots, n_s. \quad (1)$$

For each of the GNN and LAP methods we set the first column of \mathbf{P} to \mathbf{I} .

For each $i \leq n_t$ we establish a track with single-point initialization by setting

$$\begin{aligned} x_{i1} &= y_{i1}, \\ \sigma_{i1}^2 &= \sigma^2 \end{aligned} \quad (2)$$

3.3. Recursive Computation

For each s such that $2 \leq s \leq n_s$.

For each $i \leq n_t$.

For each $j \leq n_o$.

$$\begin{aligned} w &= y_{js} - x_{is} \\ \sigma_{ji}^2 &= \sigma^2 + \sigma_{is}^2 \end{aligned} \quad (3)$$

$$L_{ij} = \left[\ln(2\pi\sigma_{ji}^2) + w^2/\sigma_{ji}^2 \right] / 2$$

The likelihood matrix, \mathbf{L}_s on scan s , $2 \leq s \leq n_s$, is the matrix $\{L_{ij}\}_{1 \leq i \leq n_t, 1 \leq j \leq n_o}$.

The assignment solution vector is \mathbf{J}_s and the associated cost is

$\mathbf{C}_s, s = 1, 2, \dots, n_s$.

The scan s states, $x_{is}, i = 1, 2, \dots, n_t; 2 \leq s \leq n_s$ and variances are recursively updated according to

$$\begin{aligned}
 z &= x_{i(s-1)} \\
 \tau_{is}^2 &= \sigma_{i(s-1)}^2 \\
 K &= \sigma^2 \left(\tau_{is}^2 + \sigma^2 \right)^{-1} \\
 r &= y_{J_s(i)s} - z \\
 x_{is} &= z + Kr \\
 \sigma_{is}^2 &= (1 - K) \tau_{is}^2, \quad i = 1, 2, \dots, n_t
 \end{aligned}
 \tag{4}$$

P is updated by setting the s^{th} column of P to J_s .

After n_s scans method Purity, P , is computed as:

$$\begin{aligned}
 P &= \left[\sum_{s=2}^{n_s} \sum_{i=1}^{n_t} p(i, s) \right] / (n_s n_t), \\
 p_{is} &= 1 \text{ if } P(i, s) = i, \\
 p_{is} &= 0 \text{ if } P(i, s) \neq i.
 \end{aligned}
 \tag{5}$$

Track state accuracy, E , is computed as:

$$E = \sqrt{\left[\sum_{i=1}^{n_t} (x(i, n_s) - \mu(i))^2 \right] / n_t}
 \tag{6}$$

4. Results

We ran three 100 Monte-Carlo 10 scan MATLAB simulations characterized by the number of tracks (objects) and decreasing values of σ . The MATLAB code was run on an HP 2.2 GHZ computer. The figures in **Tables 1-5** are the average over the 100 samples. The figures in **Table 6** and **Table 7** are respective GNN and LAP assignment solution costs and on the last of the 100 samples for Case 3, $\sigma = 0.25$.

Case 1: $n_t = n_o = 3$

Table 1. Comparative accuracy performance.

σ	Purity	
	GNN	LAP
1.0	0.416	0.600
0.5	0.437	0.871
0.25	0.424	0.997
σ	Accuracy	
	GNN	LAP
1.0	0.940	0.604
0.5	0.871	0.227
0.25	0.880	0.081

Case 2: $n_i = n_o = 5$

Table 2. Comparative purity performance

σ	Purity	
	GNN	LAP
1.0	0.305	0.485
0.5	0.322	0.771
0.25	0.346	0.997

Table 3. Comparative accuracy performance.

σ	Accuracy	
	GNN	LAP
1.0	1.560	0.802
0.5	1.542	0.332
0.25	1.525	0.076

Case 3: $n_i = n_o = 10$

Table 4. Comparative purity performance

σ	Purity	
	GNN	LAP
1.0	0.206	0.454
0.5	0.230	0.780
0.25	0.269	0.990

Table 5. Comparative accuracy performance.

σ	Accuracy	
	GNN	LAP
1.0	3.39	0.807
0.5	3.26	0.338
0.25	3.26	0.084

Table 6. Case 3 comparative cost.

Scan	GNN	LAP
2	7.497	7.497
3	15.320	1.333
4	8.976	1.836
5	30.356	3.493

Continued

6	34.680	3.792
7	27.892	1.489
8	49.460	2.127
9	31.545	-3.089
10	32.554	3.335

Table 7. Case 3 Comparative Z_s^2 .

Scan	GNN	LAP
2	5.488	5.488
3	17.738	13.723
4	12.850	3.958
5	133.650	3.022
6	24.214	4.745
7	31.298	17.841
8	180.884	14.128
9	222.726	13.406
10	107.723	6.597

As for comparative assignment problem solution timing, we note that GNN requires $n_i^2 + (n-1)^2 + \dots + 2^2 = n_i(n+1)(2n_i+1)/6 - 1 = O(n_i^3)$ floating point comparisons. This is the same complexity as the LAP JV algorithm but with, as seen from the tables above, far inferior purity and accuracy. We have timed the two methods on Case 3 samples. Typical GNN operating time/per problem/scan is 0.015 sec.; the LAP method solution times did not register above 0 sec.

5. Conclusions

We first observe that the LAP observation-to-track assignment solution is the maximally likely assignment solution among all the permutations on n_i letters. It is obvious from the results above that the GNN solution ranks low in minimizing assignment problem cost among the Case 3 $n_i! = 3628880$ possible assignments and is far from optimal. It is also interesting to note that GNN purity and accuracy performance essentially do not improve as σ decreases while LAP performance increases sharply as σ decreases.

Moreover, **Table 4** confirms our two previous experiences with GNN performance, *i.e.*, purity performances in the 20 percentiles. While we have relied on a simple idealized model, the results here and in [1], which exploited a more sophisticated Multi Hypothesis Tracking (MHT) data association algorithm, are mutually supportive.

Table 5 tells us that GNN error is more than an order of magnitude larger than

LAP error. We can also infer from **Table 4** and **Table 5** that high track purity and accuracy are inextricably linked. High purity is necessary for state estimation accuracy; state estimation accuracy, based on first physical principles-based trajectory dynamics for the sensor-detected objects, is necessary for high purity to accurately predict where object observations will be in the space of measurements on the next scan.

Table 6 and **Table 7** are taken from taking the last of the Case 3 100 Monte Carlo samples as a typical result.

Table 6 emphasizes the assignment cost superiority of LAP over the GNN solutions. Except for the first row entries, where the GNN and LAP algorithms both produced exact assignment solutions, we see that GNN costs are many times higher than LAP costs.

Given LAP perfect purity, as was the case with this particular sample, the statistics $Z_s^2 = 2c_s - n_i \left[\ln(2\pi\sigma^2/s) \right]$, $s \geq 2$, appearing in the third column of **Table 7** is χ^2 distributed with n_i degrees-of-freedom. The expected value of such a random variable is $n_i = 10$ because, for each $i \leq n_i$ and $s \geq 2$, the residual $r = y_{J_s(i),i} - x_{is}$, appearing in Equation (4) is $N(0, \sigma^2(s/(s-1)))$ distributed.

It is also clear that when it is not practical, in a real satellite infrared or radar sensor tracking system, to drive sensor detection measurement line-of-sight and/or range variances down beyond certain physical limits an MHT K-best hypotheses algorithm is required.

Finally, we come to the inescapable conclusion that the use of GNN for data association in tracking systems is highly problematic. It is reasonable to assume that GNN performance will degrade when dynamics and measurement equations, as functions of time evolving states of several dimensions and sensor-to-observed object geometry, are far more complicated than in our simple model. The far more likely inference is that GNN data association real world performance will be significantly worse than is shown above.

Conflicts of Interest

The author declares no conflicts of interest.

References

- [1] Danchick, R. and Newnam, G.E. (2006) Reformulating Reid's MHT Method with Generalised Murty K-Best Ranked Linear Assignment Algorithm. *IEEE Proceedings—Radar, Sonar and Navigation*, **153**, 13-22. <https://doi.org/10.1049/ip-rsn:20050041>
- [2] Jonker, R. and Volgenant, A. (1987) A Shortest Augmenting Path Algorithm for Dense and Sparse Linear Assignment Problems. *Computing*, **38**, 325-340. <https://doi.org/10.1007/bf02278710>
- [3] Murty, K.G. (1968) Letter to the Editor—An Algorithm for Ranking All the Assignments in Order of Increasing Cost. *Operations Research*, **16**, 682-687. <https://doi.org/10.1287/opre.16.3.682>
- [4] Han, Z., Wang, F. and Li, Z. (2020) Research on Nearest Neighbor Data Association

Algorithm Based on Target “Dynamic” Monitoring Model. 2020 *IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, Chongqing, 12-14 June 2020, 665-668.

<https://doi.org/10.1109/itnec48623.2020.9085030>

- [5] Konstantinova, P., Udvariev, A. and Semerdjiev, T. (2003) A Study of a Target Tracking Algorithm Using Global Nearest Neighbor Approach. *Proceedings of the 4th International Conference on Computer Systems and Technologies E-Learning*, Rousse Bulgaria, 19-20 June 2003, 290-295. <https://doi.org/10.1145/973620.973668>
- [6] Sinha, A., Ding, Z., Kirubarajan, T. and Farooq, M. (2012) Track Quality Based Multitarget Tracking Approach for Global Nearest-Neighbor Association. *IEEE Transactions on Aerospace and Electronic Systems*, **48**, 1179-1191. <https://doi.org/10.1109/taes.2012.6178056>
- [7] Shi, M., Ling, Z. and Zu, J. (2003) Association Using Modified Global Nearest Neighbor in the Presence of Bias. *Proceedings of the Chinese Control Conference*, Xi'an, 26-28 July 2003, 4688-4691.
- [8] Fukunaga, K. and Flick, T.E. (1984) An Optimal Global Nearest Neighbor Metric. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**, 314-318. <https://doi.org/10.1109/tpami.1984.4767523>
- [9] Bahata, N. and Vandana, A. (2010) Survey of Nearest Neighbor Techniques. *International Journal of Computer Science and Information Security*, **8**, 302-305.