A Bayesian Alternative to the Statistical Distance Approach to Association

C. M. Rogers
QinetiQ Ltd

Abstract The association of a measurement or measurements with an existing set of measurements is a common problem in many engineering systems. Traditionally the method of solving this problem has been based on statistical distance. However, this approach, whilst simple, implies association unless proven otherwise. An alternative approach which does not make this implication is presented. This approach makes use of Bayesian probability theory. It is concluded that the limitations of the statistical distance methods can be overcome using the Bayesian approach, but at the cost of extra complexity in many situations.

1. Introduction

The association of a measurement or measurements to an existing set of measurements is a common problem in many engineering systems. The problem can be split into two groups, the first dealing with one new measurement being associated, and the second with many new measurements being associated. A common occurrence of this problem is in tracking systems such as air traffic control, where a new measurement of an aeroplane's state (position, heading, velocity) has to be associated with an existing track of the aeroplane, from many possible tracks, or used to start a new track.

Another occurrence of this problem is found in imaging systems on production lines, where objects are selected and sorted by colour. Due to reflections, changing light conditions and environmental conditions the colour of the object seen by the camera is not constant, and so the colour measurement has to be associated with a database of potentially many colours so that the object may be sorted correctly.

There are many methods for determining associations, ranging from probabilistic methods to heuristic methods. In this paper the focus will be on probabilistic methods, as heuristic methods such as neural networks rely on training data, and go beyond the scope of a short paper. Furthermore, data sets will now be referred to as measurements, but it should be kept in mind that the principles can be generalised.

Two main probabilistic methods exist for determining associations, these being statistical distance methods and Bayesian methods.

2. 1-to-Many Association

Where only one new measurement is being associated, it may be termed a 1-to-many association problem (figure 1). In the air traffic control tracking system the association is between a new measurement and an existing track from many potential tracks. In the production line imaging system the association is between a new colour measurement and a colour from a set of possible colours from a data base. There is essentially no limitation on what form the data to be associated may take, except that, for the association to be feasible, the data in one set implies something about the data in the other set. For example, the track contains information about where the aircraft should be and in what direction it should be heading with what velocity with a certain degree of (un)certainty. This in turn implies what sort of value the new measurement of the aircraft's position should have. Alternatively, the measured velocity of the aircraft implies whether it is a jet-propelled aircraft or a propeller driven aircraft. It would then be possible to associate the velocity measurement of the aircraft with the classification of the aircraft, from many possible classifications.

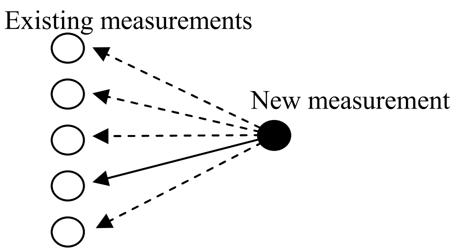

Figure 1 1-to-many association, with 1 true association (solid) and 4 alternative associations (dashed)

2.1 Statistical Distance

An uncertainty is associated with each measurement, and is expressed as the covariance of that measurement. By dividing the square of the difference of the measurements by the sum of the covariances of the measurements a dimensionless statistical distance is calculated[1].

$$D_s = (X_N - X_E)^T (Q_N + Q_E)^{-1} (X_N - X_E) \tag{1}$$

The association is then between the new measurement and that existing measurement which gives the smallest statistical distance, where the statistical distance is used rather than the physical distance as it allows for the uncertainty in the measurements. Furthermore, a threshold on the statistical distance may be chosen, such that those measurements within that threshold are deemed to be candidates for association, and those outside are deemed not to be candidates. This is effectively gating the measurements. The threshold can be chosen from the χ^2 (Chi-squared) distribution to specify the probability of truly associating measurements being accepted as associating, whilst maximising the probability that non-associating measurements will be rejected[2]. If more than one existing measurement is deemed to associate with the new measurement, then the one with the smallest statistical distance can be accepted.

The advantage of this method is that it is very simple and quick to implement. The statistical distance for every possible association between the new measurement and the existing measurements is calculated, and then a statistical distance threshold may be specified below which the measurements are deemed to associate. Furthermore, by setting a hard threshold on the statistical distance it is possible to conclude that a particular measurement does not associate with any existing measurements. In this manner, an unknown number of targets can be coped with.

The disadvantage of this method lies in the implied association assumption that it makes. Consider the case when the uncertainty in the measurements is very large. This equates to very large covariances. With such large covariances, the statistical distance between the two measurements at any one time could be small, and the measurements would therefore associate. Depending on the application, this assumed "associate unless proven otherwise" can be a severe disadvantage. In such a situation, it would be preferable to assume non-association somehow, or signify that it isn't really known whether the measurements associate or not. Furthermore, setting a hard threshold for non-association is troublesome, as defining the most appropriate value is difficult. Finally, the probability of association is not directly available using this method, which is often a desired parameter.

2.2 Bayesian Probability

The Bayesian probability method does away with the assumed association of the statistical distance method[3]. Bayes theorem states

$$P(a \mid b) = \frac{P(a)P(b \mid a)}{P(b)} \tag{2}$$

i.e. the probability of *a* given *b* is true is the *a-priori* probability of *a* multiplied by the probability of *b* given *a* divided by the unconditional probability of *b*. Applied to the association problem

$$P(\lambda_{E,N} \mid X_E, X_N) = \frac{P(\lambda_{E,N})\, p(X_E, X_N \mid \lambda_{E,N})}{p(X_E, X_N)} \tag{3}$$

where $\lambda_{E,N}$ is the association of existing measurement X_E with new measurement X_N. For the simple 1-to-many case, $P(\lambda_{E,N})$ is the inverse of the number of existing measurements, discounting the possibility of the new measurement not associating with any of the existing measurements.

$p(X_E,X_N|\lambda_{E,N})$, the probability of the data given that they associate, is simply the probability of X_N from a distribution of mean X_E and covariance Q_E+Q_N. If the distributions are assumed Gaussian,

$$p(X_E, X_N \mid \lambda_{E,N}) = \frac{e^{-\frac{1}{2}(X_N - X_E)^T (Q_E + Q_N)^{-1} (X_N - X_E)}}{\sqrt{\det\left(2\pi (Q_E + Q_N)\right)}} \tag{4}$$

$p(X_E,X_N)$, the unconditional probability of the data, cannot be calculated directly, but can be treated as a normalising factor, as if all possible associations are considered, the sum of the probabilities must equal 1, and $p(X_E,X_N)$ is constant over all $\lambda_{E,N}$.

It must be noted that if the number of targets isn't known, or the existing measurements do not cover all the targets, then assumptions must be made about the distribution of the likely number of true targets, and their spatial distribution. This is not a simple matter due to important subtleties that can lead to biasing towards association or non-association, and will not be considered in this short paper.

The probabilities of the different associations can then be used to decide which are the correct associations in the same manner as with the statistical distance method, and give the confidence of that association.

The advantages of the Bayesian method in 1-to-many association are therefore its mathematical simplicity and the fact that association is not assumed. Furthermore, a probability of association is derived, which is of great use when decision making is considered. However, there is the issue of extra assumptions required when the true number of targets is unknown.

3 Many-to-Many

The many to many association problem is significantly more complicated than the 1-to-many association problem. In this situation there are several existing measurements and also several new measurements. This means that there is not only the possibility that 1 new measurement will associate with several existing measurements, but also the possibility that 1 existing measurement will associate with several new measurements (figure 2).

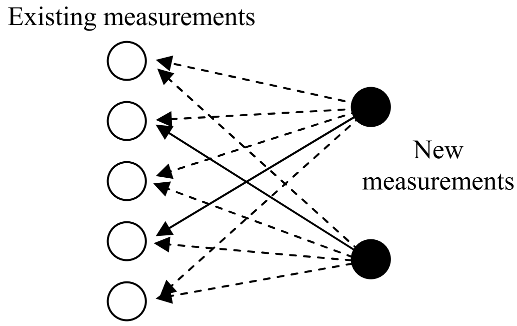

Figure 2 The many-to-many association problem

An extension to the 1-to-many statistical distance method is commonly referred to as the global nearest neighbour method, which deals with this many to many association problem. However, the disadvantages of the statistical distance method remain.

To overcome the disadvantages of the Global Nearest Neighbour method, an extension to the Bayesian method may be used. If the 1-to-many Bayesian method is considered again, each possible association, $\lambda_{N,E}$, can be termed a hypothesis, and so with 3 existing measurements and 1 new measurement, there are 4 hypotheses, including the non-association hypothesis (0), $\lambda_{N,0}$, $\lambda_{N,1}$, $\lambda_{N,2}$ and $\lambda_{N,3}$. These account for all possibilities. Similarly, when multiple new measurements are present, a set of hypotheses can be generated. For the case of 2 existing measurements and 2 new measurements, there are 7 hypotheses, including those where one or more new measurements don't associate with existing measurements. Similarly, for the case of 3 existing measurements and 2 new measurements, there are 13 hypotheses.

	N_1	N_2
$h_0 = (\lambda_{1,0}, \lambda_{2,0})$	-	-
$h_1 = (\lambda_{1,1}, \lambda_{2,0})$	E_1	-
$h_2 = (\lambda_{1,2}, \lambda_{2,0})$	E_2	-
$h_3 = (\lambda_{1,0}, \lambda_{2,1})$	-	E_1
$h_4 = (\lambda_{1,0}, \lambda_{2,2})$	-	E_2
$h_5 = (\lambda_{1,1}, \lambda_{2,2})$	E_1	E_2
$h_6 = (\lambda_{1,2}, \lambda_{2,1})$	E_2	E_1

N=2, E=2, 7 hypotheses

	N_1	N_2
$h_0 = (\lambda_{1,0}, \lambda_{2,0})$	-	-
$h_1 = (\lambda_{1,1}, \lambda_{2,0})$	E_1	-
$h_2 = (\lambda_{1,2}, \lambda_{2,0})$	E_2	-
$h_3 = (\lambda_{1,3}, \lambda_{2,0})$	E_3	
$h_4 = (\lambda_{1,0}, \lambda_{2,1})$	-	E_1
$h_5 = (\lambda_{1,0}, \lambda_{2,2})$	-	E_2
$h_6 = (\lambda_{1,0}, \lambda_{2,3})$	-	E_3
$h_7 = (\lambda_{1,1}, \lambda_{2,2})$	E_1	E_2
$h_8 = (\lambda_{1,2}, \lambda_{2,3})$	E_2	E_3
$h_9 = (\lambda_{1,3}, \lambda_{2,1})$	E_3	E_1
$h_{10} = (\lambda_{1,1}, \lambda_{2,3})$	E_1	E_3
$h_{11} = (\lambda_{1,2}, \lambda_{2,1})$	E_2	E_1
$h_{12} = (\lambda_{1,3}, \lambda_{2,2})$	E_3	E_2

N=2, E=3, 13 hypotheses

The Bayesian equation then becomes

$$P(h_i \mid X_{n1}, X_{n2}, \ldots, X_{nN}, X_{e1}, X_{e2}, \ldots, X_{eE}) = \frac{P(h_i)\, p(X_{n1}, X_{n2}, \ldots, X_{nN}, X_{e1}, X_{e2}, \ldots, X_{eE} \mid h_i)}{p\left(X_{n1}, X_{n2}, \ldots, X_{nN}, X_{e1}, X_{e2}, \ldots, X_{eE}\right)} \quad (5)$$

The *a-priori* probability of a hypothesis is constant for a constant number of new measurements, existing measurements and associations and can be simply calculated. The number of hypotheses is given by

$$\text{No\_hypes}(N,E) = \sum_{L=0}^{\min(E,N)} {}^{E}C_{L}\, {}^{N}P_{L} \tag{6}$$

where E is the number of existing measurements, N is number of new measurements and L is the number of associations in a hypothesis.

$p(X_{n1}, X_{n2},\ldots X_{nN}, X_{e1}, X_{e2},\ldots X_{eE})$ is once again a normalising factor as it is constant for all h_i, and can be calculated once all hypotheses have been considered.

It remains to calculate $p(X_{n1}, X_{n2},\ldots X_{nN}, X_{e1}, X_{e2},\ldots X_{eE}|h_i)$. This can be calculated as for the 1-to-many case, with the individual measurements being formed together into vectors of measurements, and the covariances formed into one big covariance matrix, taking care of the order in the vectors implied by the association hypothesis. For example, for hypothesis 10

$$\mathbf{X_N} = \begin{bmatrix} X_{n1} \\ X_{n2} \end{bmatrix} \quad \mathbf{X_E} = \begin{bmatrix} X_{e1} \\ X_{e3} \end{bmatrix} \quad \mathbf{Q_{NE}} = \begin{bmatrix} Q_{n1}+Q_{e1} & 0 \\ 0 & Q_{n2}+Q_{e3} \end{bmatrix}$$

giving

$$P(X_{n1}, X_{n2},\ldots, X_{nN}, X_{e1}, X_{e2},\ldots, X_{eE} \mid h_{10}) = \frac{e^{-\frac{1}{2}(\mathbf{X_N}-\mathbf{X_E})^T \mathbf{Q_{NE}^{-1}} (\mathbf{X_N}-\mathbf{X_E})}}{\sqrt{\det(2\pi\ \mathbf{Q_{NE}})}} \tag{7}$$

Through these calculations, it is possible to calculate the probabilities of individual associations by accumulating the probabilities from those hypotheses that contain the particular association of interest (figure 3).

	N_1	N_2	$P(h_i \mid \mathbf{X_E},\mathbf{X_N})$%
$h_0 = (\lambda_{1,0}, \lambda_{2,0})$	-	-	2
$h_1 = (\lambda_{1,1}, \lambda_{2,0})$	E_1	-	0
$h_2 = (\lambda_{1,2}, \lambda_{2,0})$	E_2	-	40
$h_3 = (\lambda_{1,3}, \lambda_{2,0})$	E_3		2
$h_4 = (\lambda_{1,0}, \lambda_{2,1})$	-	E_1	12
$h_5 = (\lambda_{1,0}, \lambda_{2,2})$	-	E_2	1
$h_6 = (\lambda_{1,0}, \lambda_{2,3})$	-	E_3	10
$h_7 = (\lambda_{1,1}, \lambda_{2,2})$	E_1	E_2	1
$h_8 = (\lambda_{1,2}, \lambda_{2,3})$	E_2	E_3	20
$h_9 = (\lambda_{1,3}, \lambda_{2,1})$	E_3	E_1	1
$h_{10} = (\lambda_{1,1}, \lambda_{2,3})$	E_1	E_3	1
$h_{11} = (\lambda_{1,2}, \lambda_{2,1})$	E_2	E_1	10
$h_{12} = (\lambda_{1,3}, \lambda_{2,2})$	E_3	E_2	0
			$\Sigma\, P(h_i \mid \mathbf{X_E},\mathbf{X_N}) = 100\%$

	E_1	E_2	E_3
N_1	2	70	3
N_2	23	2	31

Association Probabilities (%)

Figure 3 Calculation of association probabilities from hypothesis probabilities

4 Conclusion

Association of measurements to existing measurements is of key importance in many engineering systems. Two probabilistic methods have been presented, the statistical distance method, and the Bayesian probability theory method. Of these the statistical distance method has the severe disadvantage of assuming association unless proven otherwise, and does not provide a probability of the association being correct. This is overcome using the Bayesian method, which makes no assumptions about association and does provide a probability of association. However, the Bayesian method is more complex, especially when the number of targets is unknown.

References

[1] Bar-Shalom Y, Li X, Multitarget-Multisensor Tracking: Principles and Techniques, ISBN 0-9648312-0-1, 1995

[2] Zhang Y, Leung H, Lo T, Litva J, Distributed sequential nearest neighbour multitarget tracking algorithm, IEE Proc. – Radar, Sonar Navig., Vol. 143, No. 4, August 1996

[3] Whitehead J, Data Fusion techniques for multispectral seekers – final report, QinetiQ internal report, 2002

The author is an EngD student at the Dept. of Electronic and Electrical Engineering, University College London

The work described in this paper has been funded by the EGC Domain of the MOD CRP.